DOCUMENT RESUME

ED 055 986 SP 005 349

AUTHOR	Persell, Caroline Hodges
TITLE	The Quality of Research on Education; An Empirical Study of Researchers and Their Work. Final Report.
INSTITUTION	Columbia Univ., New York, N.Y. Bureau of Applied Social Research.
SPONS AGENCY	Office of Program Planning and Evaluation (DHEW/OE), Washington, D.C.
BUREAU NO	BR-8-0884
PUB DATE	May 71
GRANT	OEG-08-080884-4507(010)
NOTE	422p.
EDRS PRICE	MF-$0.65 HC-$16.45
DESCRIPTORS	*Behavioral Sciences; *Data Analysis; *Educational Research; *Research Methodology

ABSTRACT

This study measured and explained the quality of recent research on education. A total of 1100 articles and research papers published in 1967-68 were used as a data source and were qualitatively analyzed with respect to their substance and methodology. A stratified random sample of 390 papers was evaluated by 39 judges for contribution to theory, contribution to educational practice, and use to research methods. A 12-page questionnaire was mailed to all the authors to collect data on their background, training, career lines, attitudes, research context, and topic selection. All the data was analyzed using multivariate tubular analysis, and authors trained in schools of education were compared with those trained in the behavioral sciences. Professional socialization was found to be strongly related to research quality, and behavioral scientists did better research than education doctorates. The positive relationship between research socialization and quality was contingent upon certain background characteristics, with female, areligious, and Jewish subjects doing better research than males and those belonging to other religious groups, and young researchers doing better than older ones. Other related factors were institutional setting and being a staff member of a university research center. (Author/MBM)

ED055986

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Final Report

Project No. 8-0884

Contract No. OEG-08-080884-4507(010)

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
~~Caroline Hodges~~
Persell
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

THE QUALITY OF RESEARCH ON EDUCATION:

AN EMPIRICAL STUDY OF RESEARCHERS AND THEIR WORK

Caroline Hodges Persell
Bureau of Applied Social Research
Columbia University
New York, New York 10025

May 1971

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Copyright © 1971 by Caroline Hodges Persell

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
Office of Program Planning and Evaluation

SP005349

ABSTRACT

THE QUALITY OF RESEARCH ON EDUCATION: AN EMPIRICAL STUDY OF RESEARCHERS AND THEIR WORK

Caroline Hodges Persell

The problem of this study is one of measuring and explaining the quality of recent research on education. The problem was approached by locating all empirical research on education published in 113 American behavioral science and education journals in 1967-68 (833 articles) plus all empirical research papers presented at the 1968 annual meeting of the American Educational Research Association (an additional 267 papers). Three types of data were collected on this sample of 1,100 papers and their authors:

(1) All 1,100 research papers were qualitatively analyzed with respect to their substance and methodology.

(2) A stratified random sample of 390 papers was evaluated by a national panel of 39 judges with respect to three dimensions of quality--contribution to theory, contribution to educational practice, and use of research methods. A five point scale, developed and validated in earlier research by the author, was used to rate each of the three dimensions of quality. A sample of papers was rated in common by five judges each, to assess agreement among the judges. Consensus was measured by the average deviation from the mean, and the average of the average deviations was .62. Papers eliciting the greatest and least agreement among judges

were qualitatively analyzed in an effort to ascertain what is related to greater or lesser consensus in evaluation. Characteristics of judges were also analyzed to see if they were related to variation in agreement.

(3) A twelve page written questionnaire was mailed to all 1,100 authors in the sample, and 83 per cent returned completed questionnaires. The questionnaire collected data on the authors' background, training, career lines, attitudes, research contexts, and topic selection.

These three types of data were analyzed using multivariate tabular analysis. Throughout the analysis, authors trained in schools of education were compared with those trained in the behavioral sciences.

Professional socialization (as measured by type of doctorate, practice-oriented socialization, graduate research experience, career research experience, and research orientation) was found to be strongly related to research quality, and to have a stronger relationship to quality than undergraduate academic performance. Behavioral scientists had more research-related socialization and did better research than education doctorates.

The positive relationship between research socialization and quality was contingent upon, or intensified by, certain background characteristics. In particular, females and those who classified themselves as either areligious or Jewish were more likely than males or those belonging to other religious groups to do better research when they had more research socialization. This finding was partially explained by the greater intellectual orientation of the females, Jews, and areligious researchers in the sample studied. In addition, younger researchers were generally more likely than older ones to do better research, in part because of their greater research socialization.

Other experiences besides research socialization were related to better quality, including institutional setting (working in a university or in a private research agency rather than in a college, school system, or state department of education); and being a staff member of a university research center.

In addition to the foregoing analyses, the study explores the relationship between research quality and subsequent rewards, specifically the rewards of being promoted in academic rank, moving from a less prestigious to a more prestigious department, and applying for and receiving a research grant. The data suggest that in the behavioral sciences authors of better research papers are more likely to be rewarded than authors of average or poor papers, whereas among those trained in education, authors of better papers are no more likely to be rewarded than authors of mediocre or poor papers. Several explanations of this finding were considered.

ACKNOWLEDGMENTS

Many people have helped us in carrying out the research reported in this volume. To all the individuals listed on page 3 we owe special thanks for their assistance, stimulation, advice, and time. Unfortunately because there are so many we are unable to single out each one for the appreciation they deserve. But we can at least indicate the type of contributions each group made.

The dissertation advisors were helpful at every stage of the study from the first planning through the final writing. They gave unstintingly of their time and ideas, and their assistance is greatly appreciated. The special advisor also provided assistance and ideas throughout the study. The pre-testers made an invaluable contribution to the development of the questionnaire and they added materially to its clarity and usefulness. Our Bureau colleagues provided assistance with questionnaire construction, pre-testing, locating pertinent literature, computer programming, and the many other problems that arose while doing this research. The research assistants helped perform the many tasks--often tedious--that are involved in any collection of original data. In addition, their questions, suggestions, and ideas contributed positively to the progress of the study. Much essential assistance was pleasantly provided by many Bureau of Applied Social Research staff members. Charlotte Fisher deserves special thanks for typing the final draft. The sample validators independently scanned journals and decided whether or not articles met the criteria for inclusion in the sample.

We wish to thank the many respondents who cooperated with the study by completing the questionnaires. We are also indebted to the thirty-nine judges who read and rated eleven research articles apiece. Finally, we acknowledge with appreciation the support of the U.S. Office of Education which made this study possible.

DISSERTATION ADVISORS

Paul F. Lazarsfeld
Sam D. Sieber

SPECIAL ADVISOR

Cora Beebe, USOE

PRE-TESTERS

Roger Anderson
Samuel Ball
Miriam L. Goldberg
JB Lon Hefferlin
Elena Ruiz-Dias
Marvin Sontag
William Spady
Doxey Wilkerson
Richard Wolf

BUREAU COLLEAGUES

Richard Alba
Camilla Auger
David Caplovitz
Deborah David
Richard Faust
Nathalie Friedman
Charles Kadushin
Paula Kleinman
Bernard Levenson
Karen Seashore Louis
Adair Lummis
Anne Mahoney
Raymond Maurice
Margaret Nelson
Theresa Rogers
Lois Sanders
Wagner Thielens
David Wilder
John Ziesel

RESEARCH ASSISTANTS

Karen Crocker
Laura Falbo
Gerald Geddes
David Helfant
Rebecca Zames Margulies
Peter Shaw
James Thompson

DATA PROCESSING AND CLERICAL STAFF

Ray Albury
Takas Arsenis
Judy Barton
Ayele Berhane
Pearila Brickner
Rosa DeWitt
Elizabeth Elton
Charlotte Fisher
Nelson Glover
Lilly Grab
Pnina Grinberg
Helen Hudouskova
Liivi Joe
Evelyn Ledyard
Fred Meier
Gloria Nabors
Constance Pratt
Phyllis Sheridan
Madeline Simonson
Angela Sinanian
Pedro Soto
Sanford Wilkes

SAMPLE VALIDATORS

Sophie Fein
Nancy Walbek
Norman Walbek

TABLE OF CONTENTS

Chapter

I. INTRODUCTION 13
II. THE NATURE OF THE POPULATION: A DESCRIPTION OF THE PAPERS AND RESEARCHERS 20

PART ONE: DESCRIBING AND MEASURING RESEARCH QUALITY

III. DESCRIBING AND MEASURING RESEARCH QUALITY 55
IV. ADMINISTRATION OF THE RATING FORM AND AGREEMENT AMONG THE JUDGES 79
V. THE DISTRIBUTION OF RESEARCH QUALITY IN THE POPULATION AND THE RELATION OF THE DIMENSIONS OF RESEARCH QUALITY TO EACH OTHER 113

PART TWO: EXPLAINING RESEARCH QUALITY

VI. ROLE SOCIALIZATION AND QUALITY 152
VII. BACKGROUND AND QUALITY 199
VIII. RESEARCH SETTINGS AND QUALITY 243
IX. THE REWARD AND COMMUNICATION SYSTEMS 270
X. SUMMARY AND CONCLUSIONS 301
GLOSSARY 311
BIBLIOGRAPHY 318

APPENDIXES

II-1 JOURNALS CONSIDERED BUT ELIMINATED WITHOUT SCANNING . . 333

II-2 JOURNALS SCANNED BUT NOT YIELDING ARTICLES 340

II-3 JOURNALS SCANNED AND YIELDING ARTICLES 354

II-4 COVER LETTER TO AUTHORS ACCOMPANYING QUESTIONNAIRE . . 360

II-5 QUESTIONNAIRE FOR AUTHORS OF RESEARCH ON EDUCATIONAL TOPICS 361

II-6 ONE-PAGE QUESTIONNAIRE 374

II-7 CATEGORIZATION OF JOURNALS BY TYPE 375

IV-1 DETAILED CRITERIA RATING FORM 378

IV-2 LETTER TO JUDGES 389

IV-3 INSTRUCTION LETTER TO JUDGES 390

IV-4 DISTRIBUTION OF JUDGES' RATINGS BY JUDGE 391

VI-1 SUPPLEMENTARY TABLES FOR CHAPTER VI 395

VI-2 MATRIX OF CORRELATIONS OF ITEMS IN INDEX OF PRACTICE-ORIENTED SOCIALIZATION AND DESCRIPTION OF THE INDEX 398

VI-3 INDEX OF DEPARTMENTAL PRESTIGE 401

VI-4 MATRIX OF CORRELATIONS OF ITEMS IN CAREER RESEARCH EXPERIENCE INDEX AND DESCRIPTION OF THE INDEX 405

VII-1 SUPPLEMENTARY TABLES FOR CHAPTER VII 409

VII-2 INTELLECTUAL ORIENTATION BY RELIGION 411

VIII-1 INTELLECTUAL ORIENTATION BY ORGANIZATIONAL AFFILIATION 412

IX-1 INDEX OF PRODUCTIVITY 413

IX-2 RESEARCH QUALITY BY RANK BY AGE 414

IX-3 INDEX OF MOBILITY IN DEPARTMENTAL PRESTIGE 415

IX-4 INDEX OF BEING REWARDED 416

LIST OF TABLES

Table		Page
II-1	Percentage of Papers from Different Types of Journals	25
II-2	Primary Substance of Educational Research Papers	29
II-3	Primary Methodology of Research Papers in the Population	32
II-4	Main Methodology by Main Substance	34
II-5	Religion of Researchers in the Population	36
II-6	Percentage Researchers of Different Ages	40
II-7	Research Productivity in This Population	41
II-8	Percentage of Papers Done as Student Work	42
II-9	Regional Distribution of Researchers in Study and Public School Enrollment (Fall 1966)	43
II-10	Organizational Affiliation of Researchers	44
II-11	School or Department of Researchers in Colleges or Universityes	45
II-12	Current Specialties of Researchers	46
II-13	Total Cost of Project on Which Research was Based	47
II-14	Funded Research by Cost of Project	48
II-15	Cost of Research by Student Status by Being Funded . . .	49
II-16	Percentage of Projects Funded by Different Sources . . .	50
II-17	If U.S. Government, What Agency Funded Project	50
III-1	Percentage of Judges Remembering Authors of Papers of Different Quality	64

Table Page

IV-1 A Comparison of Two Different Administrations of an Earlier Global Rating Form 83

IV-2 Average Deviations from the Mean Rating of Nine Research Papers Rated in Common 90

IV-3 Judges' Ratings of Paper 876, the One with Highest Overall Agreement among Raters 92

IV-4 Quality and Agreement among Judges 94

IV-5 Judges' Ratings of Paper 1018, the One with Least Overall Agreement among Raters 97

IV-6 Judges' Ratings of Paper 124, the One with Highest Agreement on Research Methods 106

IV-7 Judges' Ratings of Paper 886, the One with Least Agreement on Research Methods 108

IV-8 Comparison of Deviations on Three Dimensions and on Mean Score 109

IV-9 Judges' Ratings of Paper 886 110

V-1 Distribution of Ratings of Contribution to Theory 113

V-2 Judges' Evaluations of Why Papers are Poor on Substantive Contribution 115

V-3 Summary of Judges' Comments about Theory 115

V-4 Wandt's Ratings Related to Theory 116

V-5 Distribution of Ratings of Contribution to Practice . . . 117

V-6 Distribution of Ratings of Use of Research Methods . . . 118

V-7 Judges' Reasons for Rating a Paper Poor on Research Methods 119

V-8 Summary of Judges' Comments on Shortcomings in Data Analysis 120

V-9 Summary of Judges' Comments on Shortcomings in Research Design and Data Collection 120

Table		Page
V-10	Summary of Other Comments by Judges about Research Methods	121
V-11	Wandt's Findings of Shortcomings most Frequently Cited by Experts in Evaluating 125 Educational Research Articles Published in 1962	122
V-12	A Comparison of the Ratings on the Three Quality Dimensions	123
V-13	Relationship between Theory and Practice Rating	124
V-14	Relationship between Theory and Methods Ratings	132
V-15	Relationship between Practice and Methods Ratings	140
V-16	Relationship of All Three Dimensions of the Rating Form	148
VI-1	Doctoral Major by Whether Doctorate was in Education or Not	156
VI-2	Research Quality by Type of Doctorate	159
VI-3	Research Quality by School Teaching Experience	160
VI-4	Research Quality by Name of Degree	162
VI-5	Research Quality by Number of Professional Education Courses	163
VI-6	Research Quality by Practice-Oriented Socialization	166
VI-7	Research Quality by Graduate Research Experience	169
VI-8	Research Quality by Graduate Research Experience by Doctoral Department Prestige	171
VI-9	Research Quality by Percentage of Time Spent on Research at Time Paper was Written	173
VI-10	Research Quality by Spending Six Months or More when Research was the Primary Activity	175
VI-11	Research Quality by Average Rank of Research in Author's Career	176

Table Page

VI-12 Research Quality by Staff Membership in Research Center 178

VI-13 Research Quality by Career Research Experience 180

VI-14 Research Quality by Orientation 182

VI-15 Orientation by Career Research Experience by Practice-Oriented Socialization by Type of Doctorate 184

VI-16 Orientation by Graduate Research Experience by Career Research Experience 187

VI-17 Career Research Experience by Graduate Research Experience by Type of Doctorate 190

VI-18 Research Quality by Practice-Oriented Socialization by Orientation 192

VI-19 Research Quality by Career Research Experience by Orientation 194

VI-20 Research Quality by Type of Doctorate by Orientation 196

VII-1 Research Quality by Sex 204

VII-2 Research Quality by Sex by Type of Doctorate 205

VII-3 Research Quality by Sex by Intellectual Orientation 208

VII-4 Research Quality by Religious Classification 210

VII-5 Research Quality by Religious Classification by Age 212

VII-6 Research Quality by Religious Classification by Undergraduate Performance 213

VII-7 Research Quality by Religious Classification by Type of Doctorate 215

VII-8 Research Quality by Religious Classification by Practice-Oriented Socialization 216

Table Page

VII-9 Research Quality by Religious Classification by Career Research Experience 218

VII-10 Research Quality by Religious Classification by Orientation 220

VII-11 Research Quality by Religious Classification by Intellectual Orientation 225

VII-12 Research Quality by SES (Education of Father) 229

VII-13 Research Quality by Age 231

VII-14 Research Quality by Age by Type of Doctorate 232

VII-15 Orientation by Age by Type of Doctorate 233

VII-16 Research Quality by Age by Orientation 234

VII-17 Research Quality by Undergraduate Performance 236

VII-18 Type of Doctorate by Undergraduate Performance 237

VII-19 Research Quality by (Role) Orientation by Intellectual Orientation 240

VIII-1 Research Quality by Organizational Affiliation 247

VIII-2 Orientation by Organizational Affiliation 250

VIII-3 Research Quality by Organizational Affiliation by Orientation 251

VIII-4 Research Quality by Departmental Affiliation by Type of Doctorate 253

VIII-5 Research Quality by Prestige of Department Where Wrote Paper 256

VIII-6 Research Quality by Departmental Prestige by Departmental Affiliation 257

VIII-7 Research Quality by Research Bureau Staff Status by Organizational Affiliation 260

VIII-8 Research Quality by University Research Bureau Staff Status by Intellectual Orientation 262

Table	Page
VIII-9 Research Quality by University Research Center Staff Status by Orientation	263
VIII-10 Research Quality by University Research Center Staff Status by Departmental Prestige	265
IX-1 Research Quality by Productivity (Average Annual Number of Articles)	273
IX-2 Research Quality by Productivity by Type of Doctorate	274
IX-3 Applied for and Received a Research Grant by Quality by Type of Doctorate	276
IX-4 Research Quality by Student or Faculty Status by Type of Doctorate	279
IX-5 Research Quality by Academic Rank by Type of Doctorate	281
IX-6 Change in Academic Rank by Quality by Type of Doctorate	283
IX-7 Mobility in Departmental Prestige by Quality by Type of Doctorate	285
IX-8 Research Orientation by Being Rewarded by Research Quality	289
IX-9 Applied for and Received a Research Grant by Productivity by Type of Doctorate	291
IX-10 Mobility in Departmental Prestige by Productivity by Type of Doctorate	293
IX-11 Change in Academic Rank by Productivity by Type of Doctorate	294

15

LIST OF FIGURES

Figure		Page
IV-1	Rating Form	81
IV-2	Global Rating Form	82
IV-3	Specification Form	87
VI-1	Hypothesized Model of How the Elements of Role Socialization Affect Research Quality	183
VI-2	Modified Model of How Role Socialization Affects Quality	198
VII-1	Interrelation of Background, Socialization, and Research Quality	241
VIII-1	Documented Interrelationships between Intellectual Orientation, Type of Doctorate, Research Orientation, Research Settings, and Quality	269

CHAPTER I

INTRODUCTION[1]

In this volume we are concerned with the problem of what affects performance of the research role in education, or more specifically, what processes explain variation in role performance. Our concern differs in at least two ways from that of earlier work which has touched upon this problem. First, we will assess role performance in terms of quality rather than quantity. Second, we view the processes affecting role performance as interdependent aspects of a social system. Thus we are interested in how such characteristics of the social system of research as recruitment, role socialization, career lines, orientations, organizational contexts, and the reward and communication sub-systems interact to produce research of varying quality.

We will study performance of the research role in education. Education presents a strategic research site for studying variation in the quality of research inasmuch as researchers in education are drawn from a wide array of academic fields, thereby giving us the unique opportunity to study behavioral science research in a variety of disciplines. Further, the "conjunctive domain"[1] of education, as its relation to the disciplines has been characterized, provides an opportunity to compare research in the context of a professional school with that in traditional academic disciplines, thus offering a chance to explore possible constraints imposed upon research in a professional setting.

---

[1] This research was supported in large part by a U.S. Office of Education Contract. This support is gratefully acknowledged.

In addition to serving as a site for theoretically oriented research, education is an area of great social concern. American education faces severe problems. There are the problems of educating the disadvantaged and the gifted, the difficulties of preparing children for ever increasing technological change, and the problems of coping with conflicting values and interests among different segments of society. Although these and other grave problems need attention, educational systems are so busy with their daily tasks that they are unable to generate the new understanding and techniques needed to solve such problems. Therefore, educational practitioners must look to researchers for the insight and ideas to overcome their problems and meet society's expectations. In this framework, the problem of how research on education can be improved looms large. In particular, it is assumed that if educational research is improved it will eventually benefit practice. The better the research, the more acute our understanding of educational processes, and eventually one hopes the greater the chances of solving educational problems.

---

[1] Norman W. Storer, "The Organization and Differentiation of the Scientific Community: Basic Disciplines, Applied Research and Conjunctive Domains," prepared for the Colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968 (xerox).

Our approach is first to discuss and measure the quality of research on education, and then to analyze the processes that affect it. Following a description of our sample in Chapter II, the volume is divided into two major parts. Part One, comprising Chapters III through V, addresses the problem of describing and measuring research quality. Part Two, comprising Chapters VI through IX, attempts to identify the major variables that are related to quality, to formulate a model of how these variables might be interrelated and finally to test the hypothesized interrelationships. Because our data reveal complex interrelationships, we have presented the model as it develops in a series of flow charts, one or more of which appear in each of the chapters of Part Two.

As just noted, Chapter II is primarily descriptive of the papers and authors in the sample studied, and of our methods of data collection. Where did the papers appear? What were the primary substantive topics of the papers and what research strategies were used? Is there any relation between the research methods and the types of problems that were studied? What is the background, training, and organizational contexts of these authors of recent research on education? These are the types of questions that will be considered.

Chapter III, the first chapter in Part One, deals with the nature of research quality and the problems of measuring it. We review several studies that have measured research outcomes, and try to specify the assumptions underlying the various methods. We consider whether some methods are more appropriate than others under certain circumstances, and present the rationale for the method used in this study.

In Chapter IV we discuss the rating form used to measure quality in this study and describe how it was administered. Then we examine the extent and nature of agreement and disagreement among the judges on commonly-rated papers, and try to ascertain whether certain characteristics of judges or papers contribute to greater or lesser consensus. In particular, we do a qualitative analysis of the papers which elicited the greatest and the least agreement among raters in an effort to understand why some research reports are rated similarly and others are not. Also, we examine judges' specialties, locations, training, age, and sex to see if the characteristics of judges explain agreement. Then we focus on each dimension of quality (theory, practice, and research methods) to see if one dimension elicits more consensus among judges than another. As a way of learning more about what affects agreement among judges on each dimension, we also qualitatively analyze the papers with the most and the least agreement among judges on each of the three dimensions. In short, the purpose of the analysis is to see if we can identify general characteristics of research papers or judges that are related to consensus among evaluators.

We present the results of the judges' ratings in Chapter V, showing the distribution of research quality in this sample of authors. Is research quality distributed normally or is it skewed toward one end of the rating scale? Then, using qualitative comments by the judges, we summarize the weaknesses in papers rated poor on substance, research design, data collection, or data analysis. In this chapter we also examine the interrelation of the three dimensions of quality, and present qualitative analyses of those papers that were rated high on one dimension and low on another. These analyses were done to provide insights into both the rating process and the nature of good research.

Part Two of this volume includes Chapters VI to IX. There we shift our focus from the phenomenon to be explained, that is, research quality, to a consideration of the processes that affect it.

One of the key processes that we believe affects research quality is role socialization, and we devote Chapter VI to a discussion of this process. Earlier empirical studies of research training have found several experiences which seem to have positive or negative consequences for research involvement or productivity. These studies, together with more general literature on socialization for professional roles, enable us to identify five elements of role socialization that we consider central for explaining variation in quality. We discuss each of these elements and then develop a model of the interrelationships among them and quality. In the second half of the chapter we analyze the relationships hypothesized in the model and try to specify the processes underlying these associations. The chapter on role socialization precedes the one on background because we considered it necessary to examine the elements of this complex phenomenon before we could explore the part recruitment plays in affecting the relationship between role socialization and quality.

In Chapter VII we turn to the question of selective recruitment. There we discuss the background of authors of educational research, including their sex, religion, socio-economic status, age, and undergraduate academic performance. Traditionally, educational research has been done by upwardly mobile, older, male Protestants. We consider the evidence and opinions about how these background characteristics may be associated with quality. We also explore whether background characteristics affect the relationship between role socialization and quality, and if so, why.

In Chapter VIII we turn from the authors' recruitment and role socialization to their work settings. Observers of educational research often mention structural settings as key factors in explaining variation in research quality. We consider several frequently mentioned settings, including: organizational affiliation, i.e., university, college, specialized agency, school system or other setting; departmental affiliation; departmental prestige; and university research center affiliation. Where relevant, we examine the interrelation of work settings, background, role socialization, and quality, in an effort to specify the conditions under which work settings are related to better research.

We move beyond the authors' immediate work context in Chapter IX to the larger social system in which he pursues his career. Here we seek to describe the reward and communication systems and to analyse how their operation may affect quality. Does the reward system encourage better or poorer research? What are some of the other consequences of the way it operates? Why does the reward system operate the way it does in this social system? These are some of the questions we consider in Chapter IX.

In Chapter X we summarize the findings of our study, and conclude by presenting a set of conditions that seem to enhance the quality of empirical research on education.

CHAPTER II

THE NATURE OF THE POPULATION: A DESCRIPTION OF THE PAPERS AND RESEARCHERS

Nature of the Population

The concern of this study is recent empirical research on education, as conducted by behavioral scientists and by people in education. It is important to focus upon discrete research papers in order to assess their quality. Therefore, we selected a time sample of recent authors of research papers whose articles were printed in education and behavioral science journals published in the United States during the period July 1967 through June 1968, or whose papers were presented at the 1968 American Educational Research Association (AERA) annual meeting.

The educational and behavioral science journals were selected as follows: First, we checked a number of bibliographic references[1] to develop an inclusive list of journals published in the United States which might contain empirical research articles on education. We deliberately included as many discipline journals as possible so we could compare education doctorates with behavioral scientists.

---

[1] We drew the education journals from the following sources: America's Education Press, a classified list of periodicals issued in the U.S. and Canada, The Educational Press Association of America, Yearbook 29, 1966; Bibliographic Index, a cumulative bibliography of bibliographies (New York: H. W. Wilson Co., 1967 and 1968); M. Blaug, Economics of Education: A Selected Annotated Bibliography (New York and Oxford: Pergamon Press, 1966); Education Index (New York: H. W. Wilson Co., Vol. 40, No. 1, 4); Saul Herner, Janet D. Griffith, and Mary Herner, Study of Periodicals and Serials in Education, Final Report, Project No. BR 7-9003, U.S. Office of Education, Department of Health, Education and Welfare (Washington, D.C.: Herner and Company, June 28, 1968), ERIC document number ED 017 747; and Psychological

Second, we excluded several classes of publication prior to scanning. These include: review journals which merely summarize existing research; publications of state teachers' associations, state departments of education, or proceedings of professional associations; dissertation abstracts; yearbooks; monographs; journals with a circulation of less than 1,000;[1] privately commissioned and circulated papers and reports (such as those done by the RAND Corporation) if they are

---

Abstracts (Washington, D.C.: American Psychological Association, 1967).

We drew the behavioral science journals from the following references: Bibliographic Index, op. cit.; Economic Abstracts, V. IV-1, No. 21, February 1956 (New York: New York University, 1956); Index of Economic Journals, Vol. VII, 1964-1965, prepared under the auspices of the American Economic Association (Homewood, Ill.: Richard D. Irwin, Inc., 1967); International Bibliography of the Social Sciences: Economics, 1966, Vol. XV (Chicago: Aldine; London: Tavistock, 1968); International Bibliography of the Social Sciences: Political Science, Vol. XV (Chicago: Aldine; London: Tavistock, 1966); International Bibliography of the Social Sciences: Social and Cultural Anthropology, Vol. XII (Chicago: Aldine; London: Tavistock, prepared under the auspices of the International Union of Anthropological and Ethnological Sciences by ICSSD, with the support of UNESCO, 1967); International Bibliography of the Social Sciences: Sociology (Chicago: Aldine; London: Tavistock, 1966); International Political Science Abstracts, Vol. XVII, No. 1, October-December 1966 (Oxford: Basil Blackwell, 1967); Psychological Abstracts, op. cit.; and Sociological Abstracts (New York: International Sociological Association, Eastern Sociological Society, and Midwest Sociological Society, Vol. 13, Nos. 1-7, covering 1965, c. 1964-67).

[1] We determined circulation of less than 1,000 by consulting the following: America's Education Press, a classified list of periodicals issued in the U.S. and Canada, Yearbook 29 (Educational Press Association of America, 1966); and Directory: Newspapers and Periodicals 1968 (Philadelphia, Pa.: N. W. Ayer and Sons, 1968).

not reported in journal articles; and journals dealing with physical education, home economics, industrial arts, nursing, or school safety.

A list of 503 journals emerged from the above selection process. Of these, we eliminated 131 without scanning because they were beyond the defined scope of the population, they were impossible to locate, or for some other reasons such as not publishing empirical research, or being a regional, review, or foreign publication. (See Appendix II-1.) We scanned the remaining 372 journals for the designated year. Of these journals 259 did not yield any empirical research articles about education (Appendix II-2 contains the names and reasons). The balance, or 113 journals (Appendix II-3), yielded a total of 846 empirical research articles on education.[1]

In addition to the 846 journal articles, the sample includes 269 empirical research papers on education presented at the 1968 annual American Educational Research Association (AERA) meeting. The people in our sample, then, are recent authors of empirical research papers on education, which was either published in a journal or presented at the 1968 AERA meeting. We included only those authors with addresses in the United States because the study was confined to American educational research.[2] If an author had published more than one paper during the time period, we randomly eliminated all but one of his papers.[1] That paper was the one we mentioned in the cover letter and on the first page of the questionnaire sent to all 1,100 authors in the sample. In the case of papers with several authors, we included the first-named author of the article in the sample.

---

[1] A sample of articles was scanned independently by three other people from several fields, as a way of verifying our decisions about which articles should be included in the population. We agreed in all but six cases, which were borderline topics dealing with fairly basic research studies in psychology. After considering the situation, we decided to include these borderline cases.

[2] In addition to the six articles mentioned in note 1, we subsequently discovered twelve people, such as foreign students, who had

One indicator of the nature of this sample is its inclusion of "significant" researchers.[2] When we have informally asked people in education what they thought were come of the most significant recent research studies in education, they inevitably mention the Coleman report[3] and the Rosenthal and Jacobson study.[4] Both of these studies are represented by papers in the population. Also, other prominent researchers including Urie Bronfenbrenner, John B. Carroll, John L.

---

returned or moved to another country. They were deleted from the final sample. In addition, another three were eliminated for the following reasons: two were found to be in the population twice (under different, married names), so one of them was randomly eliminated; the judges deemed the papers of two others not to be empirical research, and we felt that one paper did not deal with education. After these changes the total number of authors in the sample was 1,100.

[1] This was done so no researcher would receive more than one questionnaire. Also, since the others were eliminated randomly, the one rated can be assumed to be representative. Fewer than 10 per cent of the researchers had more than one paper in the population.

[2] One member of the U.S. Office of Education questionnaire clearance unit asked why we were not studying authors of research monographs, since that was where much important research was being done. Our answer was that this broad a population of researchers would include many, if not all, of the types of researchers he had in mind.

[3] James S. Coleman, et al., Equality of Educational Opportunity (Washington, D.C.: Government Printing Office, 1966).

[4] Robert Rosenthal and Lenore Jacobson, Pygmalion in the Classroom (New York: Holt, Rinehart, and Winston, Inc., 1968).

Holland, Robert C. Nichols, Thomas F. Pettigrew, William Sewell, Julian C. Stanley, Patrick Suppes, and Donald Super, among others, are in the sample. As a result, we feel that the way we defined the sample was useful for the purposes of this study.

Types of Data Collected

We mailed a twelve-page printed questionnaire accompanied by a cover letter (Appendix II-4) on May 22 and 23, 1969 to the 1,100 authors in the sample. At the cut-off date of November 24, 1969, 82 per cent of the sample had returned the questionnaire (Appendix II-5), and another 3 per cent had returned a one-page abbreviated questionnaire (Appendix II-6) sent to all non-respondents who were not in the rating sample. The long questionnaire provides data on the background, training, career line, research involvement, orientations, and work contexts of researchers in the sample.

We found no significant differences between respondents and non-respondents on number of co-authors, region, location, field, type of journal in which they published, main substance and methodology of their paper, and number of other papers they had in the sample. The biggest difference between respondents and non-respondents was that between AERA paper-givers and people publishing journal articles. Only 73 per cent of the AERA paper-givers responded, compared to 85 per cent of those publishing articles. This difference probably occurred because AERA paper-givers had already been surveyed for several other studies before we approached them. A smaller difference between respondents and non-respondents occurred on sex. Males were 6 per cent more likely to respond than females.

In addition to the data collected about authors, we had a random sample of 390 research papers rated by a national panel of 39 judges. (See Chapter IV for a discussion of how the judges were selected.) This operation provided data on the quality of educational research papers. Besides having the papers rated, we classified them according to their primary substance and methodology (design, data collection, and analysis). We turn now to a description of recent educational research in terms of this classification scheme.

Recent Educational Research

Type of Journals in Which the Papers Appeared

As we have seen in the first section of this chapter, the final sample of 1,100 consisted of 833 papers which appeared in journals and 267 which were presented at the 1968 AERA meeting. Not surprisingly, most of those published in journals appeared in educational journals.

TABLE II-1

PERCENTAGE OF PAPERS FROM DIFFERENT TYPES OF JOURNALS

Type of Journal[a]				N
Education	Behavioral Science	General	Other	
69%	26	1	4	833

[a] For a list of how the journals were classified, see Appendix II- .

Two-thirds of the articles appeared in educational journals, and a quarter were published in discipline journals. One of the differences between papers appearing in the different types of journals is their length. Articles in behavioral science journals tend to be longer than articles in education journals. The average length of articles in education journals is seven pages, compared with ten pages in behavioral science journals. Articles in general journals and other journals are also longer, their average length being thirteen pages and nine pages, respectively. For all types of journals, the articles range in length from one to forty-eight pages. Thirty-five per cent of the articles are less than six pages long, 82 per cent are less than eleven pages long, and 90 per cent are less than sixteen pages long. The average length of papers in the population is eight pages.

Primary Substance of the Research Papers

In describing the substance of the papers, we are including AERA papers with journal articles. As we would expect from the range of journals scanned and the diverse topics of AERA papers, this research includes a variety of substantive areas. In classifying papers according to their primary substance, we used the following outline as a guide:[1]

---

[1] Initially we classified all 1,100 papers in the population on their substance and methodology. Over a year later, the research assistant on this project at the time, David Helfant, again classified the journal articles in the rating sample on substance and research design and I again independently classified this sample. As a result of this triple classification, we reclassified 27 per cent of the papers as to their substance. The purpose of this repetition was primarily to refine the classifications used rather than to verify the

PSYCHOLOGICAL PROCESSES

- Learning (motivation, curiosity, cognition and perception)
- Child development not directly related to learning (self-images, personality development, values, personal adjustment, leadership)

CURRICULUM

- Methods of instruction
- Programmed instruction
- Teaching machines
- Content, quality

READING

SPEECH AND NORMAL SKILLS

- Speech, audiology
- Normal skills (hearing, sight)

SPECIAL EDUCATION

- Education of the deaf and blind
- Mentally retarded
- Talented, gifted, high-ability students
- Adult education

GUIDANCE AND COUNSELING

TRAINING OF TEACHERS

- Education as a profession
- Training

ADMINISTRATION

- School finance--allocation of non-human resources
- Educational administration and organization (other than finance)

---

reliability of the classification scheme. The purpose of classifying the papers was to provide a crude basis for describing recent research. The major problems in classification arose with papers which covered several substantive areas, e.g., higher education and guidance and counseling. We tried to ascertain the primary area, and classify it accordingly.

SOCIAL CONTEXT OF EDUCATION

- Sub-cultural differences
- Other social problems: drop-outs, delinquency, aged
- School-community relations
- Comparative education
- Career patterns and occupational choice

HIGHER EDUCATION

RESEARCH METHODS

- Tests and Measurements
- Other research methodology
- Organization and conduct of education research

These are broad groupings, but they help provide some feeling for the nature of recent educational research. Nearly one in five of the research papers in this sample deals with some aspect of psychological processes--learning, personality, development (Table II-2). This is not particularly surprising since education has long been characterized by a psychological approach. Many of the pioneering contributors to research on education were psychologists, such as Thorndike, Hall, and Terman. Some observers suggest that educational research has been dominated by psychology. For example, Arthur Foshay the research coordinator at Teachers College of Columbia University, states:

> Inquiry in education, having begun as social bookkeeping, proceeded after the turn of the century to come under the domination of the new field of psychology. It has not made use of other disciplines for the formulation of its research problems with anything like the skill available in those fields; it has been limited to what the psychologists could see.[1]

---

[1] Arthur W. Foshay, "Issues and Dilemmas in Nurturing the Educational Researcher in an Organizational Setting," in The Training and Nurture of Educational Researchers, p. 172.

TABLE II-2

PRIMARY SUBSTANCE OF EDUCATIONAL RESEARCH PAPERS
(N=1,100)

	%
Psychological processes	19%
Curriculum	18
Higher education	12
Personnel teacher training	11
Guidance and counseling	10
Research methods	9
Social context	8
Reading	5
Special education	3
Administration	2
Speech	2
Other	1
	100%

Another notable finding in Table II-2 is the relatively large amount of research on higher education. One explanation is that higher education is a different type of category than the others and often includes them, particularly in the areas of guidance and social context. Where the focus of such papers was a problem in higher education, then we coded it as being higher education rather than one of the other substantive areas. Even with the possible inflation of this category by classification procedures, however, there has been an increase in research on higher education over the last few years. Compare Sieber and Lazarsfeld's[1] classification of 160 research papers published in 1964, where higher education was not even included as a substantive area. The increase in research in this area is undoubtedly related to the enormous growth of higher education in the last ten years. The percentage of research in each of the other substantive areas in Table II-2 is similar to Sieber and Lazarsfeld's findings.

Research Methodology of the Papers

In addition to classifying papers according to their primary substance, we also grouped them according to the type of research design employed. We defined the following types of research design:

Experiment: A true experiment is an attempt to make inferences of causation or association with the aid of controlled exposure to stimuli, either through artificially structuring the situation or through artful selection and arrangement of the subjects in real life. The experimental stimulus must be presented to an aggregate of people randomly selected from a common population (called an experimental group) and not presented to another randomly selected aggregate (control group). Then an observation is made to determine the differential effect of the presentation of the stimulus.[2]

Quasi-experiment: The term "quasi-experiment" refers to situations where the researcher can introduce something like experimental design into his scheduling of data collection procedures (e.g., the _when_ and _to whom_ of measurement), even though he lacks the full control over the scheduling of experimental stimuli (the _when_ and _to whom_ of exposure and the ability to randomize exposures) which makes a true experiment possible.[1]

---

[1] Sieber and Lazarsfeld, The Organization of Educational Research, Appendix D-11.

[2] Written by Marshall Childs of International Business Machines, after the definition in Donald Campbell and Julian Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally & Co., 1963).

Sub-group comparison: A broad classification for the methodological approach that Campbell and Stanley include under the heading of "correlational and _ex post facto_ designs."[2] For example, deaf and hearing children are compared on certain psychological tests. The researcher is interested in learning more about the difference between deaf and hearing children, such as how they compare on tests which are used to measure certain characteristics. Another example is where children scoring high and low on a particular test are grouped, and then certain correlations with other variables or test results are run to learn more about what is associated with high and low scores. Such correlations do not necessarily show causation, but they can indicate the presence of new factors which may have been unknown before, and thus pave the way for future research.

We have not included definitions of the other methods because we feel they are self-explanatory.

Since 31 per cent of the papers use either the experimental or the quasi-experimental method (Table II-3), it seems that psychology has influenced more than just the selection of research topics in

---

[1] _Ibid._, p. 34.

[2] _Ibid._, p. 64.

TABLE II-3

PRIMARY METHODOLOGY OF RESEARCH PAPERS IN THE POPULATION (N=1,100)

	Percentage	
Sub-group comparisons	22%	
Quasi-experiment	18	31%
Experiment	13	
Statistical analyses[a]	10	
Analytic survey	8	
Descriptive survey	8	
Panel or trend study	6	
Factor analysis[a]	5	
All other research methods[b]	11	

[a] While these are methods of analysis rather than strategies of data collection, certain studies focused primarily on data analysis, and therefore we coded them accordingly.

[b] Including secondary analysis, census data analysis, content analysis, field observation.

education. Research designs common to psychology are used even in papers that do not deal with psychological processes. After experiments and quasi-experiments, sub-group comparison is the research desi design most frequently used (22 per cent). Sub-group comparisons may be the easiest design to use since they do not require the controls that an experiment or quasi-experiment does, nor do they require the large-scale sampling and data processing of analytic surveys. The remainder of the studies are fairly evenly divided in their use of other research methods, with no method employed in more than 10 per cent of the papers.[1]

Interrelation of Substance and Methodology

We expected to find a strong relationship between the substance and methodology of these papers because the type of methods used limits the types of variables considered, and the variables determine what research questions can be raised. Substance and methodology are strongly related (Table II-4). The upper left-hand quadrant of Table II-4 (in the broken lines) illustrates how frequently experiments and quasi-experiments are used in studies of curriculum, reading, and speech, as well as in studies of psychological processes.

Similarly, analytic and descriptive surveys, which are associated with the field of sociology, are frequently used to study the social context of education or problems in higher education and administration. (See the section within the dotted line in the right hand

---

[1] As a result of the triple classification of papers on their research methodology, we re-classified 33 per cent of the papers on research methods.

TABLE II-4

MAIN METHODOLOGY BY MAIN SUBSTANCE

Main Methodology	Psychological Processes	Curriculum	Reading	Main Substance: Speech	Main Substance: Special Education	Guidance & Counseling	Personnel & Teacher Training	Research Methods	Adminis-tration	Higher Education	Social Context	Other	
Experiment	49	53	13	2	3	9	9	1	2	3	1	2	147
Quasi-experiment	48	71	14	10	5	15	19	9	1	5	4	1	199
Descriptive survey	4	10	4	1	4	14	20	1	2	20	5	0	85
Analytic survey	4	8	0	0	0	11	8	1	7	28	22	0	89
Sub-group comparisons	63	22	9	4	3	27	29	25	4	34	16	3	239
Panel or trend study	9	11	4	0	3	10	7	8	1	9	5	1	68
Statistical analysis	12	10	8	2	2	7	15	22	2	11	13	3	107
Factor analysis	17	5	0	2	3	4	9	6	0	1	3	0	50
All others	4	6	4	2	7	14	9	28	5	15	20	2	116
	210	196	56	23	30	111	122	101	24	126	89	12	1,100

portion of Table II-4.) Other substantive areas, particularly psychological processes, guidance, teacher training, research methods, and higher education, seem to use sub-group comparisons most frequently. (These relationships are circled on Table II-4.)

Characteristics of the Researchers in This Population

Demographic Characteristics

Having learned something about the research papers in this sample, we will now briefly describe the researchers themselves.

Sex

Not surprisingly, 85 per cent of the sample is male, 15 per cent female. This is exactly the same proportion of males and females that Barger found in the population of educational researchers he studied in 1964.[1] Perhaps because there are so few females they are more highly motivated and selected than males, and thus do better research. We will consider this possibility in Chapter VII.

Race

This sample is 97 per cent white, 2 per cent black and 1 per cent Oriental. While many social and historical factors may be found to explain the low proportion of blacks, our concern is with the effects of such underrepresentation on research. Among the crucial issues in education today is the education of urban dwellers and the nature and effects of racism in our society. Having so few black researchers who might bring their background to bear on these and related issues probably has crucial consequences for educational research. Because there are so few blacks in this population we are unable to test the hypothesis that race is related to research topic selection and quality.[1]

---

[1] Robert Barger, et al., "Development of a National Register of Educational Researchers" (Columbus, Ohio: The Ohio State University Research Foundation, 1965).

Religion

This population is more diverse in terms of religion than with regard to sex or race (Table II-5).

TABLE II-5

RELIGION OF RESEARCHERS IN THE POPULATION
(N=900)

Protestant	Areligious	Catholic	Jewish	Other	No Answer	
45%	20	12	11	7	6	=101%

The most notable finding is the large number of people indicating "none" as their religion. It is surprising compared with the population at large, but not compared to other scientists, researchers, and university people. It is also not surprising in terms of Weberian[2] theory which was related to the rise of science by Merton,[3] if we extend Weber's theory to include agnostics and atheists as well as Protestants and Catholics. His thesis, as extended by Merton, is that certain religious beliefs are more congruent with scientific inquiry than others. It is possible that the lack of religious belief, or skepticism, is the conceptual style most conducive to scientific inquiry. If so, we would expect the areligious to do better research than people with some organized religion.

---

[1] The small proportion of blacks in our sample could be due to sampling error or it could accurately reflect the small proportion in the population of educational researchers. The small number of blacks at recent AERA meetings supports the latter conclusion.

[2] Max Weber, The Protestant Ethic and the Spirit of Capitalism.

[3] Robert K. Merton, Social Theory and Social Structure (New York: Free Press of Glencoe, 1957; 2nd ed.), pp. 531-628.

Jews are also overrepresented in this sample compared with their numbers in the population at large. There are several explanations for this. Jews in general have often entered and excelled in academic life. Strodtbeck, for one, noted that Jews have higher educational aspirations than other ethnic and religious groups.[1] Also, most Jews are from urban areas and the source of talent has been shifting steadily in this century from rural to urban areas, as the population has been shifting. There are many population figures to document this, but they would not paint the picture as vividly as the personal observation of Ralph Tyler. In reflecting on a study of Ph.D.'s done by the National Academy of Sciences for the period through the 1950's which showed that the chief source of Ph.D.'s had shifted to the city colleges, he said:

> The only hypothesis that I have seen that seems to fit all the data thus far obtained from both sides is that the greatest opportunities for young people, in the days when Edgar Dale and Guy Buswell and I came out of the Plains states, were in education, if one's father wasn't wealthy and if one couldn't go into

---

[1] Fred L. Strodtbeck, "Family Interaction, Values, and Achievement," in Talent and Society, ed. by David C. McClelland, Alfred L. Baldwin, Urie Bronfenbrenner, and Fred L. Strodtbeck (New York: Van Nostrand, 1958), p. 174.

> his medical or law office or take over his business. An educator could move out of Nebraska or North Dakota and get to more comfortable and satisfactory places to work. We came in large numbers. That's no longer true. . . . The great place where you now find young people with ability who have limited opportunities is in the big cities, and this has been the chief source since World War II, according to the National Academy study.[1]

This shift affects more than Jews, of course. For a moment, however, we will consider the significance of disproportionate numbers of Jews for educational research. In his study of reading researchers, Wilder found that Jews were more oriented toward research than members of other religions. He found further that this difference was not explained by the greater tendency of Jews to receive the Ph.D. (rather than the Ed.D.) nor by their greater attendance proportionately at top ten universities. Religion remained independently related to research orientation, with type of degree and quality of school having largely independent effects on research orientation.[2] In view of Wilder's findings that Jewish reading researchers were more involved in research, we expect that Jews may be doing better research than people of other religions. If this is the case, we will try to ascertain what is associated with being Jewish that leads to doing better research or whether differences in quality are due to differences in undergraduate performance prior to graduate school, to graduate training experience, to career lines, or to some other factor.

The proportion of Protestants in this sample is less than we would expect, given the national population figures and the composition of American education. Wilder's population of reading researchers was 70 per cent Protestant, although he found that Protestants were proportionately less involved in research than the areligious or Jews. In view of Tyler's observation about the opportunity for upward mobility which education affords, it may be that Protestants now need this opportunity less.

---

[1] Ralph Tyler, in The Nurture of Educational Researchers, p. 71.

[2] David Wilder, "The Reading Experts" (unpublished Ph.D. dissertation, Columbia University, 1966), Appendix D.

Catholics are also underrepresented in this sample, as they were in Wilder's as well. In fact, Catholics have both historically[1] and recently been underrepresented among scientists.[2] As Cooper has noted, this fact has been acknowledged by Catholics in the United States.[3]

Age

When we defined this sample as people publishing in journals and giving papers at AERA, we wondered if that definition would include researchers of all ages. This time sample of researchers includes a range of ages (Table II-6). Because of the rapid growth of educational research in the last ten years, we would expect there to be more younger researchers, but older ones are also represented.

---

[1] Merton, Social Theory and Social Structure.

[2] Robert H. Knapp and H. B. Goodrich, Origins of American Scientists (Chicago: University of Chicago Press, 1952).

[3] J. M. Cooper, "Catholics and Scientific Research," Commonweal, XLII (1945), 147-49; cited in Wilder, "The Reading Experts," p. 29.

TABLE II-6

PERCENTAGE OF RESEARCHERS OF DIFFERENT AGES
(N=892)[a]

29 or Less	30-39	40-49	50-70	
7%	43	37	14	= 101%

[a] The number of cases on which tables are based varies due to non-response on certain items.

The average age in Bargar's population was 45, compared to 41 in this sample. Bargar's identification procedures may have weighted the list somewhat in favor of older researchers. To be included by Bargar, a researcher had to be already "established" in the sense that he was listed in a professional directory or journal, or recommended by a director of a state department of education, state education association, or division of research in the public schools.[1]

Another indicator besides age that we have active as well as fledgling researchers in this sample is the research productivity of the group. Half of the researchers have published six or more research papers in addition to the one on which they were selected (Table II-7). Twenty-six per cent have published one or more books. Therefore this population represents the working researcher, as well as the new one.

---

[1] Bargar, "Who is the Educational Researcher?" in The Nurture of Educational Researchers, p. 15.

TABLE II-7

RESEARCH PRODUCTIVITY IN THIS POPULATION
(N=900)

	Number of Papers Published in Addition to One in Population						
	No Others	One	Two-Three	Four-Five	Six-Ten	Over Ten	
Percentage of researchers	15%	7	16	11	18	32	= 99%

	Number of Books Published					
	None	One	Two-Three	Four-Five	Six or More	
Percentage of researchers	75	13	9	2	2	= 101%

Training Characteristics

All but 11 per cent of the 900 respondents have their doctorates, and some were in graduate school when they answered the questionnaire, so eventually more than 90 per cent of these authors will have doctorates. Of those with doctorates, 66 per cent received them in education and 34 per cent did not. This division is similar to the distribution of research articles in different types of journals.

In the population Bargar studied, 82 per cent had doctorates.[1] He did not ask whether or not their highest degree was in education, so we do not know how the proportions in his population compare with ours. In another study Buswell selected a sample of educational researchers, all of whom had doctorates from schools of education.[2]

---

[1] Ibid.

[2] Guy T. Buswell, et al., Training for Educational Research (Berkeley: University of California, 1966).

Thirty-eight per cent of these papers were done to fulfill academic requirements, mostly for the doctorate (Table II-8).

TABLE II-8

PERCENTAGE OF PAPERS DONE AS STUDENT WORK
(N=895)

	Paper was done as part of:				
	Doctorate	Master's	Other Student Paper	None of These	
Percentage of researchers	28	5	5	63	= 101%

Publishing one's student work shows an early orientation toward research. In Chapter IX we compared the quality of research done by students and faculty.

Current Location

Region

Researchers in this sample work in all regions of the United States (Table II-9). The largest proportion of researchers is currently located in the middle west, which is consistent with public school enrollment figures. In general there are no great discrepancies between the school enrollment and the regional distribution of recent researchers.

TABLE VII-9

REGIONAL DISTRIBUTION OF RESEARCHERS IN STUDY AND PUBLIC SCHOOL ENROLLMENT (FALL 1966)

Region[a]	Researchers in Study	Public School Enrollment[b]
Midwest	34%	30%
Northeast	27	23
West	20	24
South	19	23
	100%	100%
	(1,100)	(43,055,055)

---

[a] The regions were defined as follows: Midwest: Dakotas Nebraska, Kansas, Oklahoma, Missouri, Illinois, Indiana, Ohio, Michigan, Wisconsin; Northeast: Maine through Pennsylvania, Delaware, New Jersey; West: New Mexico, Colorado, Wyoming, Utah, and west and north; South: Washington, D.C., Maryland, West Virginia, Virginia, Kentucky, Arkansas, Texas, and below.

[b] Source of data is Alvin Renetzky and Phyllis Ann Kaplan, Standard Education Almanac, 1968 (Los Angeles: Academic Media, Inc., 1968).

Setting

Eighty-four per cent of the sample were working in colleges or universities when they published or presented their papers (Table II-10).

TABLE II-10

ORGANIZATIONAL AFFILIATION OF RESEARCHERS
(N=1,100)

Setting	Percentage
University	75%
College	9
School system	7
Private agency	5
Government agency	1
State department of education	.5
Other (e.g., hospital, business, etc.)	3
	100.5%

Of the 84 per cent who are in colleges or universities, almost half are in a school or department of education (Table II-11), which is less than we expected since 66 per cent received their doctorates in education.

TABLE II-11

SCHOOL OR DEPARTMENT OF RESEARCHERS IN COLLEGES OR UNIVERSITIES

	Percentage	
School or Department of Education		
Undergraduate	3%	48%
Graduate	22	
Joint undergraduate/graduate	23	
Liberal Arts Department		
Undergraduate	4	20%
Graduate	4	
Joint undergraduate/graduate	12	
Other School or Department		
Undergraduate	2	14%
Graduate	4	
Joint undergraduate/graduate	8	
No answer to this question	6	
People not in College or University	17	
	105%[a]	
	(901)[b]	

---

[a] The total is greater than 100 per cent because some of the 901 respondents were affiliated with more than one school or department.

[b] The sample size here is 901 rather than 1,100 because this information was obtained from the respondents' questionnaires. Information that was coded from the papers themselves is presented in tables with a sample size of 1,100.

Current Specialty

Given the proportion who received their doctorates in education and the percentage who are now in schools or departments of education, it is not surprising that 47 per cent indicate their current specialty as being some aspect of education (Table II-12). The proportion is even greater if educational psychology is considered part of education.

TABLE II-12

CURRENT SPECIALTIES OF RESEARCHERS

	Percentage	
Guidance and counseling	11%	47%
Subject matter area in education (e.g., art, business, language, math, music, science, etc.)	8	
Teacher training	7	
Curriculum	6	
Tests and measurements	6	
Educational administration	5	
Higher education	4	
Educational psychology	19	
Other psychology	14	
Sociology	5	
Other[a]	14	
	99%	
	(892)	

[a] Most did not specify what their specialty was, if they indicated other. We know some are in economics.

In the population Bargar studied, 52 per cent indicated that education was their specialty, 40 per cent indicated psychology, 5 per cent sociology, and the others scattered among different fields.[1]

Research Support

A surprising 63 per cent of respondents (or 558 people) report that the project on which their research paper was based received financial support from some source. Half of the projects cost between $500 and $10,000 (Table II-13).

TABLE II-13

TOTAL COST OF PROJECT ON WHICH RESEARCH WAS BASED

Amount	Percentage
Under $100	9%
$101-$499	20
$500-$10,000	50
$10,001-$50,000	10
$50,001 +	10
	99%
	(719)

Not everyone whose project cost something was funded. Of the 719 who listed cost figures for their projects, 558 were funded, leaving 155 (17 per cent) who were not funded. The more expensive the project, not surprisingly, the more likely it was to have been funded (Table II-14).

---

[1] Bargar, Development of a National Register, p. 29.

TABLE II-14

FUNDED RESEARCH BY COST OF PROJECT

	Cost					
	Under $100	$101-500	$501-10,000	$10,001-50,000	Over $50,000	
Percentage of research which was supported	19%	49	74	86	99	
N =	(65)	(146)	(361)	(73)	(73)	= 718

There is a positive linear relationship between the cost of the research project and the percentage of projects funded (Table II-14). We would expect that projects done by students would generally cost less. Also, we might expect that students would be less likely to be funded. Student projects do generally cost less than non-student work (Table II-15). A student's chances for being funded, however, are actually greater than those of non-students if the total cost of the project is between $500 and $10,000. One possible explanation of this finding may be the emphasis of the Small Grants program of the U.S. Office of Education which tries to make grants of $10,000 or less available to people who have not been funded before. This would include most students, and the Small Grants program does seek to aid dissertation projects. The single largest source of research support is the U.S. government (Table II-16) and the majority of U.S. government funds are from the U.S. Office of Education (Table II-17).

TABLE II-15

COST OF RESEARCH BY STUDENT STATUS
BY BEING FUNDED

	Was Research Funded?			
	Yes		No	
	Student Work	Not Student Work	Student Work	Not Student Work
Under $100	2%	3%	19%	28%
$101-$500	20	12	38	28
$501-$10,000	69	49	45	34
$10,001-$50,000	8	15		10
$50,000 +	1	22	—	—
	100%	101%	102%	100%
N =	(156)	(327)	(128)	(102) = 713

TABLE II-16

PERCENTAGE OF PROJECTS FUNDED BY DIFFERENT SOURCES
(N=558)

U.S. Govt.	College or Univ.	Govt. & Other	Private	State	Other[a]
42%	32	10	10	3	1

[a] For example, city or school system

TABLE II-17

IF U.S. GOVERNMENT, WHAT AGENCY FUNDED PROJECT
(N=321)[a]

Office of Education	NIMH	OEO	NSF	Other[b]	
63%	9	9	6	12	= 99%
					NA = 21

[a] This figure is larger than 42 per cent of 558 apparently because funds received by the researcher from his college or university were originally from a government agency.

[b] Includes DOD, DOL, NASA, ONR, PHS, and Children's Bureau.

We are interested in the relationship between research funding and research quality. This is a much-debated issue in the field. Some people argue that all that is needed to improve research on education is more funds. Others suggest the opposite, that a reduction of available funds is the best way to improve research. For example, "Cole suggests that funding agencies supporting research on education should cut back their budgets rather than give grants to people who don't present adequate credentials or adequate proposals."[1] Given the range of opinions on this issue, we will scrutinize the relationship between research funding and quality in Chapter IX.

Summary

In this chapter we have described how we selected a time sample of recent researchers on education. In short, our time period was the fiscal year 1967-68. We included everyone with an American address who published an empirical research article on education in one of 113 education and behavioral science journals during that time. In addition, we included everyone giving an empirical research paper on education at the 1968 AERA meeting. The final sample totaled 1,100 authors. All of them were surveyed with a mailed 12-page questionnaire to collect data on their background, training, career lines, research involvement, and attitudes. Of the 1,100, 82 per cent (or 901) returned the 12-page questionnaire and another 3 per cent returned a one-page questionnaire. When we compared respondents and non-respondents on existing data, there were no significant differences with respect to region, setting, number of co-authors, type of journal in which they published, main substance and methodology of their paper, and number of other papers in the population.

---

[1] Ronald G. Corwin and Maynard Seider, "Patterns of Educational Research: Reflections on Some General Issues" (paper prepared for the American Educational Research Association Colloquium, November 21-22, 1968). (Mimeographed.)

In addition to surveying the researchers, we had a national panel of 39 judges rate the quality of a stratified random sample of 390 papers. They used a five-point rating scale to assess the paper's contribution to theory, contribution to practice, and use of research methods.

Finally, we qualitatively analyzed the research papers in the population with respect to their primary substance and research design. Papers of particular interest were qualitatively analyzed in greater detail (see Chapters IV and V).

In this chapter we have briefly described the nature of the researchers and their papers. We have seen that 69 per cent of the articles appeared in education journals and 26 per cent in behavioral science journals. The two biggest substantive areas of papers are psychological processes and curriculum and the most frequently used research designs are experiments and quasi-experiments together, followed by sub-group comparisons. Substance and methodology tend to be related.

Turning to the researchers, we saw that 85 per cent are male and 97 per cent are white. Forty-five per cent are Protestant, 20 per cent are areligious, 12 per cent are Catholic, 11 per cent are Jewish, and 7 per cent consider themselves to be some other religion. We considered some of the implications of these findings for educational research and noted some of our expectations about possible relations between religion and research quality. The average age of these researchers is 41. Eighty-nine per cent have their doctorates, with 66 per cent in education and 34 per cent in the behavioral sciences. More researchers are from the midwest than any other region (34 per cent). The vast majority (84 per cent) are in colleges or universities and of these, almost half are in schools or departments of education. The majority (47 per cent) indicate that their current specialty is some aspect of education.

Finally, the majority (63 per cent) of these researchers based their papers on projects that were funded. These characteristics of researchers will be considered in relation to research quality in the course of this study.

PART ONE

DESCRIBING AND MEASURING RESEARCH QUALITY

CHAPTER III

DESCRIBING AND MEASURING RESEARCH QUALITY

The Nature of Research Quality

The literature in the field has treated research quality in several ways. Sometimes the quality of research is referred to very generally, assuming the reader knows what is meant by the concept. For example, Corwin and Seider state, "The underlying importance of fostering high quality research was designated by Zuckerman as the chief problem facing educational research."[1] They say nothing more about what is meant by the concept of high quality research.

Often a reference to research quality is illustrated with negative statements which reveal what the lack of research quality means. For example, Bereiter sees the failings of educational research as "a failure to make discoveries and a failure to carry out studies that add up to anything or converge upon anything."[2] Quite obviously, he is implying that good research would make discoveries and conduct studies which cumulatively contribute to important problems.

---

[1] Corwin and Seider, "Patterns of Educational Research," p. 10.

[2] Carl Bereiter, "Issues and Dilemmas in Developing Training Programs for Educational Researchers," in The Training and Nurture of Educational Researchers, Sixth Annual Phi Delta Kappa Symposium on Educational Research (Bloomington, Indiana: Phi Delta Kappa, 1965), p. 99.

Another negative statement also suggests what the author sees as research quality. Guba notes,

> Educational research has not had the impact that one might expect of it. Whether one expects research to make significant discoveries about the world, to act as a change agent, to influence teachers in the classrooms, to provide order and structure for a discipline of education, or to formulate and develop educational science, educational research can hardly be said to have fulfilled our fond hopes for it.[1]

His assessment of the field implies that better research would make a difference.

Two other negative comments deal more specifically with the nature of research itself. The philosopher of science, Scriven, states that "by minimum acceptable research standards, 95 percent of the work [on education] . . . that is concerned with causal analysis is, by either theoretical or practical standards, invalid or trivial."[2] This statement suggests that quality research meets acceptable research standards, is valid, and significant. We are not, however, apprised of what these standards are.

Buswell is the most specific in his negative appraisal of research quality:

> One of the most serious deficiencies of educational research is that it is still composed mainly of fragmentary, small-scale investigations at a time when research on human behavior is no longer produced mainly by individual scholars but is increasingly the product of collaboration. . . . Another serious limitation

---

[1] Egon G. Guba, "An Overview of the Symposium," in The Training and Nurture of Educational Researchers, Sixth Annual Phi Delta Kappa Symposium on Educational Research (Bloomington, Indiana: Phi Delta Kappa, 1965), p. 276.

[2] Michael Scriven, "The Philosophy of Science in Educational Research," Review of Educational Research, XXX (December 1960), 426.

> of current educational research is that it is of relatively unimaginative and uncomplicated design, in spite of the fact that methods of multivariate analysis have been developed during the last two decades. . . . Seldom, if ever, do differences in learning correspond to variation in a single variable.[1]

Buswell indirectly indicates that better research would consist of large-scale, integrated studies of multiple factors.

The most explicit statements about research quality come from people who have expressed the concept in terms of measurable indicators. Wandt, Nasatir, Gephart, and Persell each delineate the concept of research quality in the form of evaluation check lists.[2] They judged research quality to consist of the following aspects:[3]

A clear statement of the problem, concepts, and hypotheses.

A significant problem (W,N,P).

A statement of relationships to previous research and existing knowledge.

---

[1] Buswell, et al., Training for Educational Research, p. 1.

[2] Edwin Wandt, et al., "An Evaluation of Educational Research Published in Journals," report of the Committee on Evaluation of Research, American Educational Research Association, February, 1967. (Mimeographed.) A summary of this report appears in Edwin Wandt, A Cross-Section of Educational Research (New York: David McKay Company, Inc., 1965), pp. 1-13. David Nasatir and David Elesh, "Project Memorandum No. 5" of The Organization of Educational Research (New York: Bureau of Applied Social Research, Columbia University, 1965). (Mimeographed.) William J. Gephart, The Development of an Instrument for Evaluating Reports of Educational Research (Ann Arbor, Michigan: University Microfilms, No. 16-13, 23Q, 1965). Caroline Hodges [Persell], "Measuring Educational Research Quality and Its Correlates" (unpublished Master's essay, Columbia University, 1967).

[3] Where everyone did not include a given dimension, the initials of those who did mention it are listed in parentheses.

A full description of the design used.
Use of an appropriate design.
Lack of major weaknesses in the design (G and P provide checklists of many possible weaknesses).

Appropriate methods of analysis.
Analysis correctly done.
Analysis related to initial hypotheses (N,G,P).
Analysis presented clearly (W,N,P).

Conclusions substantiated by the data.
Conclusions that contribute to theory or practice (P).
Conclusions that raise questions for further research (N,G,P).

These efforts may be taken as a specification of the more general terms used by others discussing research quality, such as "competent," "valid," "worthwhile," "important," "significant," "having impact."

Another specification of research quality is that which appears in Cole and Cole.[1] They define quality, in effect, as significance, with citation rates as the measure of significance. This specification reveals what papers are being used most frequently, but does not reveal why certain papers are deemed significant.

It is clear from the preceding discussion that judgments of people in a field play an important role in the definition of research quality. There does not seem to be some intrinsic characteristic of quality which can be recognized by an objective instrument like a geiger counter. Even Gephart, who made a monumental effort to create an objective instrument for measuring quality, had to rely on judges to ascertain both the occurrence of the items on his checklist and to evaluate the significance of their contribution to the value or soundness of the research paper.[1] In short, even when standards are made explicit, their application requires judgments by people. This does not mean that consensus is impossible, it just means that it is a variable which needs to be understood in its own right.

---

[1] Stephen Cole and Johnathan R. Cole, "Scientific Output and Recognition: A Study in the Operation of the Reward System in Science," American Sociological Review, XXXII (June 1967), 377-90.

Methods of Measuring Research Quality

Students of scientific work have measured research output with the following indicators: citation rates, evaluation forms, research productivity, prize winning, peer evaluations, characteristics of the journal publishing a paper, and self-evaluations. Each of these criteria has certain assumptions.

Citation Rates

The advent of the Science Citation Index (SCI) in 1961 provided a convenient measure of research output. In a study of physicists, Cole and Cole used the number of citations to represent the "relative scientific significance or 'quality' of papers."[2] The Coles indicate several problems in the use of citations as a measure of quality. First, they note that scientists publishing a large number of papers each of which is cited only a few times might accumulate as many citations as those publishing a few heavily cited papers.[3] They handle this problem by considering a physicist's three most heavily cited contributions. Second, they observe that the extreme contemporaneity of science raises a problem in the use of citations as an indicator of quality. At least half the citations in a year are to work published in the five preceding years. They handle this problem by a technique of weighting citations.[1]

---

[1] Gephart, Development of an Instrument, pp. 72-73.

[2] Cole and Cole, "Scientific Output," p. 379.

[3] Ibid., p. 380.

They note other problems in the use of the SCI as well:

> (1) Work of the highest significance often becomes common knowledge very quickly and is referred to in papers without being cited.
> (2) Citations may be critical rather than positive.
> (3) The various scientific fields differ in size. If we wish to compare the work of scientists in different fields, we must take into consideration the number of people actively working in these fields.
> (4) The significance of scientific work is not always recognized by contemporaries (e.g., Mendel).[2]

The problem of not being recognized by one's contemporaries is not unique to citation rates as a measure of quality. It applies to all the measures of research output we will consider, except productivity and self-evaluation.

For citation rates to be a valid measure of research quality, we must accept the assumptions on which they rest. First, use of citation rates assumes that there is equal visibility throughout the system, i.e., an unknown researcher publishing in an obscure journal is just as likely to be cited, if his work is of comparable quality, as a well-known researcher publishing in the leading journal in the field. Under conditions of equal visibility, better research will be visible, regardless of the rank of a researcher's department, the prestige of his honorofic awards, or the nature of his specialty. (By visibility the Coles mean familiarity of other physicists with a given researcher's work.) The Coles analyze the interrelation of several variables and visibility, and conclude that quality, departmental rank, prestige of awards, and specialty all have an independent effect on visibility.[1] Thus, quality is but one of several factors related to visibility. Since work must be visible before it can be cited, unequal visibility may affect citation rates.

---

[1] Ibid.

[2] Ibid.

Equal visibility requires at least that there be an established core of journals in a given field, which virtually everyone reads or at least scane, and also that everyone has an equal chance of being published in that core of journals, assuming equal quality in their papers. In physics, these conditions are met to a considerable degree, although the principle of cumulative advantage[2] combines with universalism[3] in the operation of the referee system, as Zuckerman and Merton note. They write, "Physicists who have acquired a reputation for doing good work have certain advantages in getting their papers through the referee process. In general, their work is more apt to be accepted but we cannot say whether this is because it tends to be of higher quality or because it is more often given the benefit of the doubt."[1]

---

[1] Stephen Cole and Jonathan R. Cole, "Visibility and Structural Bases of Awareness in Science," American Sociological Review, XXIII, No. 3 (1968), 397-412.

[2] By the principle of cumulative advantage, or the "Matthew effect" as Merton calls it, he means "the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark." Robert K. Merton, "The Matthew Effect in Science," Science, January 5, 1968, pp. 56-63.

[3] By universalism is meant the uniform application of standards for rewards to everyone in the social system.

There is some evidence that visibility may vary between scientific fields and among "conjunctive domains," as well as according to the prominence of individuals in a particular field. In the field of educational research, for example, Garvey, Nelson, and Lin found that people presenting papers at the 1968 American Educational Research Association (AERA) meeting named a total of seventy-six different journals to which they either had submitted or planned to submit the main content of their presentations for publication. They found that "an educational researcher must examine eighteen different journals in order to read half the material presented at the AERA meeting. Compared to most other groups, AERA seems extraordinarily diffuse in its range of publication vehicles."[2] Garvey, Nelson, and Lin present data showing the range of publication vehicles for eleven different professional associations (see Chapter IX). Clearly we cannot assume that visibility would be identical in fields with such varying ranges of publication outlets, without first empirically verifying this assumption. (We discuss the communication system further in Chapter IX.)

---

[1] Harriet Zuckerman and Robert K. Merton, "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System" (paper presented at the annual meeting of the American Sociological Association, 1968), pp. 50-51.

[2] William D. Garvey, Carnot Nelson, and Nan Lin, "A Preliminary Description of Scientific Information Exchange in Educational Research" (paper presented at the Colloquium on Improving the Social and Communication Mechanisms of Educational Research, sponsored by the American Educational Research Association, November 21-22, 1968, Washington, D.C.).

One way to ascertain whether all research papers are equally visible would be to measure quality in a different way, e.g., with a rating form, and then ask judges if they recognize the paper or the author (when the author and journal names and other identifying characteristics have been removed). We did this with the 390 educational research papers rated by 39 judges, and found that the judges were familiar with only 8 per cent of the research papers. Not surprisingly, in view of the Coles' finding, they were more likely to be familiar with better research than with poorer research. There is a positive linear relationship between familiarity and rated quality on theory and practice. The relationship between familiarity and research methods is also positive, although not linear (Table III-1). To be sure, this is just one paper out of perhaps many which an author may have written. It is possible that other work of the author's would be visible and hence cited. Given the diffuse communication system in educational research, however, we do not want to assume that good work has an equal chance of being cited regardless of who wrote it or where it appears.

Also underlying the use of citation rates as a measure of quality is the assumption that citation rates are unaffected by location, social circles, or other social structural variables. Since we have already seen from the Coles' work that visibility is related to rank of department and prestige of honorific awards received, even when research quality is held constant, and that ease and speed of publication are related to a researcher's reputation[1] (which unquestionably is

---

[1] Zuckerman and Merton, "Patterns of Evaluation in Science."

TABLE III-1

PERCENTAGE OF JUDGES REMEMBERING AUTHORS OF PAPERS OF DIFFERENT QUALITY

	Rated Quality of Paper				
	Best	Above Average	Average	Below Average	Incompetent
	(1)	(2)	(3)	(4)	(5)
Contribution to theory[a]	25% (16)	10 (98)	2 (100)	1 (92)	2 (68)
Contribution to practice[a]	18 (22)	7 (94)	3 (128)	3 (91)	0 (41)
Use of research methods	11 (19)	11 (93)	3 (118)	1 (90)	2 (53)

[a] Both relationships have a Chi-Square which is significant above the [illegible] level.

related to quality) it is hard to assume that citation rates are unaffected by social structural variables.

Another principle that seems to underlie the use of citation rates as a measure of research quality is the assumption that there is high consensus on what the standards for assessing research quality are, and high agreement on the way these standards are applied. In other words, variation in citation rates is assumed to depend primarily on the paper's value, rather than upon uncertainty about standards for assessing the research or uncertainty about how those standards should be applied. This assumption may be valid in a well institutionalized field like physics, but not valid in a poorly institutionalized field like education, although this assumption has not been empirically tested by the users of the measure to our knowledge.

Also, using citation rates assumes that everyone making citations (i.e., publishing papers in journals) is an equally competent judge of quality. The validity of this assumption probably depends on the level of institutionalization of research in a field, and all that institutionalization implies about the existence and operation of standards in a field.

As a measure of research output, citation rates tap one dimension that other measures do not, however. If a research paper is not cited, then, no matter what its "quality" it is not being used by other researchers. Citation rates measure use as well as quality.[1]

---

[1] This aspect of citation rates was suggested by Camilla Auger, Bureau of Applied Social Research, Columbia University.

Finally, there are practical factors to consider when comparing methods of measurement. Citation rates are relatively easy and inexpensive, and they permit considerable scope in the physical sciences, and increasingly in the behavioral sciences. The relatively minor problems noted by the Coles notwithstanding, they provide a good way of ascertaining how much a researcher's work was used by his colleagues, which is certainly a key dimension of quality. They may present serious problems, however, in fields with high variations in visibility and in the application of standards of research. A practical limitation in their use is the required time lag of several years before enough citations have accumulated to be used as a measure. This means they cannot be used to measure the quality of very recent research papers.

Evaluation Forms

A major difference between citation rates and evaluation forms as measures of research quality is that the latter explicitly present the standards used in making an assessment. Also, the amount of consensus in the application of those standards can be measured. Forty-eight different evaluation forms were counted by Bartos.[1] The problems of measuring research quality with evaluation forms were thoroughly treated by Wandt, Gephart, Nasatir, and Persell. We will consider the standards they used, the consensus they found among raters, and the assumptions which must be made in using this type of measure.

---

[1] Bruce B. Bartos, "A Review of Instruments Developed to be Used in the Evaluation of the Adequacy of Reported Research" (Bloomington, Indiana: Phi Delta Kappa, 1969), p. 1. (Mimeographed.)

Wandt and his colleagues[1] developed and administered two measures of research quality. They developed a twenty-five item list of characteristics of good research, with each item being rated on a five point scale. Of their twenty-five characteristics, all but one deals with either research methods or with style of presentation. Only one criterion considers the substantive worth of the problem being treated by the article ("The problem is significant") and that does not specify significant in terms of either theoretical contribution or practical contribution. They also asked each judge to assume the role of an educational research journal editor and rate the article as one he would accept, accept with specified minor revisions, or reject. This was called the Accept-Revise-Reject Rating (ARR). Wandt had each of 41 papers rated by two judges. On the overall ARR rating the reliability coefficient was .17, which is significant at the .20 level.

Wandt recommends having judges who are experts in the substantive area of the article, rather than experts in educational research methods generally. Also, he recommends omitting the name of both journals and authors being rated by the judges, so this source of bias might be eliminated. (Both of these recommendations were followed in the study reported here.)

---

[1] Edwin Wandt, Georgia S. Adams, Dorothy M. Collett, William B. Michael, David G. Ryans, and Carleton B. Shay, "An Evaluation of Educational Research Published in Journals," report of the Committee on Evaluation of Research, American Educational Research Association, February, 1967. (Mimeographed.) A summary of this report appears in Edwin Wandt, A Cross-Section of Educational Research (New York: David McKay Company, Inc., c. 1965), pp. 1-13.

Another helpful contributor to the literature of measuring educational research quality is William J. Gephart.[1] He developed a rating instrument composed of fifty items and an overall measure which was presented in a clear outline format, facilitating use. His instrument reflects the psychological-experimental nature of much educational research. He tested the instrument's reliability with ten qualified judges who were specialists in guidance and counseling, and they rated articles from five guidance and counseling journals. The reliability of the ratings was .599 (using the average of Kendall's coefficients of concordance), which is significant at the .001 level. Unlike Wandt, Gephart found higher agreement on the overall rating of the studies than on the fifty item instrument.[2] On the overall ratings Gephart's judges revealed high agreement as indicated by their Kendall's W which was .807. (Both the Kendall's W and the Chi Square obtained on these results are significant at the .001 level.)[3]

Gephart also compared two ways of measuring quality, by rating and ranking, and found that the degree of agreement obtained when judges rate research projects and when they rank them is nearly the same. (He found Spearman's rank different coefficient correlation on the two measures to be .981, significant beyond the .01 level.) In addition, he found higher interrater reliability in ranking and rating research articles when the judge was a specialist in the same substantive area as the research article.

---

[1] Gephart, Development of an Instrument.

[2] Ibid., p. 103.

[3] Ibid.

Gephart's purpose was to evaluate the research methods used. As a result, he does not consider the importance of the problem selected for study. He hints at the dimension of importance when he says, "This assessment is based upon the premise that good research is carefully planned, quite traditionally analyzed, and adequately reported."[1] It could be, however, that very important work might initially be done and reported in a haphazard way, and only later be verified in a methodologically sound way. Also, a research article might be very sound technically, but treat a trivial subject. For these reasons, we feel it is essential to assess both substance and research methods when measuring quality.

Nasatir developed a detailed evaluation form with eleven criteria and a global rating form.[2] He had ten judges rate six articles apiece. Like Gephart, he found more agreement on the global rating form than on the detailed instrument. On the global form, 70 per cent of the judges were in the average modal frequency with respect to substance and 68 per cent were in the average modal frequency with respect to methods.

Drawing on the work of Gephart, Nasatir, and Sieber,[3] Persell developed another detailed rating instrument.[4] She tested its reliability by having eighteen judges (nine from the Bureau of Applied Social Research of Columbia University and nine from Teachers College of Columbia) rate 102 articles from 38 journals, on detailed and global rating forms. Consensus was measured by the percentage of judges in the average modal frequency. For the detailed rating instrument that measure of agreement was 61 per cent and on the global form it was 71 per cent. The coefficient of rank correlation for the 102 articles according to the two rating instruments was .81. In sum, the global instrument elicited higher consensus than the detailed one and the articles were ranked very similarly on both global and detailed instruments. The one-page global instrument was far easier to administer than the forty-six item, eleven-page, detailed instrument. The consensus obtained in the Persell administration was virtually identical to that obtained by Nasatir on the global form, and was somewhat higher on the detailed form.

---

[1] Ibid., p. 138.

[2] Nasatir and Elesh, "Project Memorandum No. 5."

[3] Sam D. Sieber, "Detailed Criteria of Evaluation" (rev. ed.; New York: Bureau of Applied Social Research, Columbia University, 1966). (Mimeographed.)

[4] [Persell], "Measuring Educational Research Quality."

Valid use of an evaluation form requires the following assumptions. First, that the relevant dimensions of quality and criteria for them can be determined, and agreed upon. As we have already noted, Gephart emphasized only research methods. He made no attempt to assess the substantive contribution a research paper made, perhaps feeling that this dimension of quality was more difficult for judges to assess in a consistent way. Of Wandt's twenty-five characteristics, all but one deals with either research methods or with style of presentation. Nasatir, in contrast, has six out of eleven criteria in his detailed rating form which deal with substantive aspects of a research paper.

Persell had four out of forty-six criteria that treat substance in her detailed rating instrument. These variations suggest that there is not complete agreement on what constitutes relevant criteria. But there may be two ways to assess agreement. First, are the same items included? Second, would researchers in a field vehemently deny that certain items were relevant? It is possible that they would agree that certain criteria were relevant, but not include them in their own rating form because they thought they would encounter problems of measurement. If this is the case, then researchers might agree to a standard such as "significance" if agreement were reached among those rating it.

One measure that Wandt, Gephart, Nasatir, and Persell all agreed upon is the use of an overall, or global rating, of a research paper's quality. Wandt used a single item rating, with a three point rating scale. He recommends, however, using a five point rating scale for the overall measure. Gephart also has a single item overall evaluation, using a five point scale, ranging from -2 to 2. Nasatir used a Global Rating Form in addition to his detailed rating form. His form had two dimensions--substantive contribution and contribution to research methods. Each of these two dimensions used a five point scale. Persell retained Nasatir's Global Form unchanged. Since these four major efforts agree on using an overall rating, that represents consensus on a general approach to the problem at least.

Using an evaluation form also assumes uniform administration by different judges. This assumes, further, that if a paper is recognized by a judge, his knowledge of the author does not influence his evaluation of the paper. As already noted, Gephart and Persell found higher agreement among judges on overall ratings than on detailed ones. Future studies should test the reliability of any evaluation form proposed for use.

Finally, using an evaluation instrument assumes that the measure is valid. As mentioned, Gephart tested the validity of his evaluation form by comparing the way the ratings ordered the papers with an independent ranking of the papers and found the two were highly related. As already noted, Persell found that the detailed evaluation form and the global rating form ranked articles very similarly. We have already mentioned that it would be useful to compare the results of measuring quality with evaluation forms and with citation rates. Problems in each of the measures might complicate an assessment of their correlation, but they should be strongly related, nonetheless. We turn now to another method which has been used to measure research output.

Research Productivity

Many researchers have used quantity of output as an indicator of research quality.[1] As the Coles observe, a high or moderate correlation between research productivity and its quality has been found in several fields of science. Also, in their own work they find a correlation between the research productivity and quality of work published by a sample of physicists (r = .72). As they note, since the correlation is not perfect there are low-producing scientists who are thought to have contributed a great deal and high-producing ones considered to have contributed relatively little.[1]

---

[1] For some examples see: Joseph Ben-David, "Scientific Productivity and Academic Organization in 19th Century Medicine," in Science and Society, ed. by Norman Kaplan (Chicago: Rand McNally, 1965), pp. 39-61; Buswell, et al., Training for Educational Research; Myron B. Coler, ed., Essays on Creativity in the Sciences (New York: New York University Press, 1963), p. xvi; Diana Crane, "Scientists at Major and Minor Universities: A Study of Productivity and Recognition," American Sociological Review, XXX (October 1965), 699-714; David A. Lingwood, "Interpersonal Communication, Scientific Productivity, and Invisible Colleges: Studies of Two Behavioral Research Areas" (paper prepared

In another study, Heiss used a combination of research quantity and peer evaluations to determine "outstanding scholars" in education.[2] She found ninety-four "high producing" researchers. When they were asked to name the five best researchers in their specialty, however, the list shrank to thirty-eight "outstanding scholars" who were both productive and doing good work. The relatively low-producing good scholar is lost by this technique, however.

---

for Colloquium on Improving the Social and Communication Mechanisms of Educational Research, sponsored by the American Educational Research Association, Washington, D.C., Nov. 21-22, 1968); Derek Price, Little Science, Big Science (New York: Columbia University Press, 1963), p. 40; Logan Wilson, The Academic Man: A Study in the Sociology of a Profession (New York: Octagon Press, 1964), p. 110; Blaine R. Worthen, "The Impact of Research Assistantship Experience on the Subsequent Development of Educational Researchers" (unpublished Ph.D. dissertation, Ohio State University, 1968).

[1] Cole and Cole, "Scientific Output," pp. 377-78, 382. They cite the following: Kenneth E. Clark, America's Psychologists: A Survey of a Growing Profession (Washington, D.C.: American Psychological Association, 1957), Chap. 3; Wayne Dennis, "The Bibliographies of Eminent Scientists," Scientific Monthly, LXXIX (September, 1954), 180-83; Bernard M. Meltzer, "The Productivity of Social Scientists," American Journal of Sociology, LV (July, 1949), 25-29; and Peggy Thomasson and Julian C. Stanley, "Exploratory Study of Productivity and 'Creativity' of Prominent Psychometricians" (unpublished paper, University of Wisconsin, 1966).

[2] Buswell, et al., Training for Educational Research.

In sum, it seems that although research productivity has a demonstrated positive relationship to research quality in the physical sciences, it would be better to consider it as a separate concept, and try to measure quality and quantity independently.

Prize Winning

In the natural sciences and increasingly in the behavioral sciences, awards are given to outstanding contributors in the field. Most would agree the awards are given for better work. If receipt of important prizes in a field is taken as a measure of research quality, one must be willing to make assumptions similar to those required in using citation rates, namely that of equal visibility throughout the system, agreement on the standards of assessing quality, agreement on the way those standards are applied, and competence of the people making the judgments, plus the absence of political considerations. Zuckerman has extensively studied Nobel laureates in science,[1] and the Coles have found that winning the Nobel prize is very highly related to being frequently cited.[2]

---

[1] Harriet Anne Zuckerman, "Nobel Laureates in the United States: A Sociological Study of Scientific Collaboration" (unpublished Ph.D. dissertation, Columbia University, 1965); "Nobel Laureates in Science: Patterns of Productivity, Collaboration, and Authorship," American Sociological Review, XXII (June, 1967), 391-403; "The Sociology of the Nobel Prizes," Scientific American, CCXVII (November, 1967), 25-33.

[2] Cole and Cole, "Scientific Output."

Peer Evaluations

A frequently used peer-group judgment is the referee process, applied to journals or grants.[1] This method requires the same assumptions as citation rates and prize winning, plus the additional assumption that invisible college membership, or other forms of collegial social circles, do not affect how a researcher evaluates his peers. Even in a field like physics, where research is highly institutionalized, Zuckerman and Merton found that the principle of universalism is modified by the principle of cumulative advantage in the refereeing process.

Sieber found that the background of proposal-reading committee members (i.e., in education or the behavioral sciences) was associated with approval rates of grant applicants (the applicants with backgrounds similar to the readers' had a higher percentage of approvals). These findings suggest that there are unresolved problems in the use of peer evaluations as measures of research quality.

Journal Characteristics

A number of journal characteristics might be used as indicators of journal quality, and thus indirectly as indicators of the quality of research papers. Such characteristics include:

---

[1] In reference to journals, see Richard H. Orr, et al., Peer Group Judgment of Scientific Merit (Philadelphia: Institute for Advancement of Medical Communication, in process); Zuckerman and Merton, "Patterns of Evaluation in Science." In reference to grants see Sieber and Lazarsfeld, Organization of Educational Research, Appendix A, "Features of Research Proposals Submitted to the Cooperative Research Program between 1956 and 1963". For other peer-evaluations, see: Clark, America's Psychologists, and Heiss, in Buswell, et al., Training for Educational Research.

(a) relative citation rates (i.e., is a journal being cited more or less often than one would expect from the proportion of articles in that journal relative to all articles in the field?);[1]

(b) a referee system of selecting papers for publication;

(c) rejection rates (although this seems to vary by field, with the humanities having the highest, followed by the behavioral sciences, and physical sciences;[2]

(d) page charges (although this varies by field);

(e) critical reviews of research papers published in a journal and rebuttals by the author;[3]

(f) readership; and

(g) expert opinion about what are considered the leading journals in the field.

Using journal characteristics as an indirect measure of research quality has several major limitations. Certain journals have special emphases, and hence they might reject good research because it was beyond their stated scope. Also, this measure of quality would not include good work published in lesser journals. It seems that too little is known at this time even about the relation of the above indicators to journal characteristics, much less about the relation of those indicators to the quality of research papers, to be able to use journal characteristics as a measure of quality.

---

[1] Mentioned in Zuckerman and Merton, "Patterns of Evaluation in Science."

[2] Ibid., p. 51.

[3] Suggested in a personal communication from William Asher, Professor of Education and Psychology, Purdue University, 1967.

Self Evaluations

For obvious reasons, self-evaluation of research quality cannot be considered in the same class as the other measures of research output. Like productivity, self-evaluation should be considered an indicator of a different concept. As such it should be measured independently and then compared with quality. There may be some theoretical reason for expecting people who do better quality research to evaluate it in a certain way, e.g., more harshly than people doing research of lower quality. Shaw studied productivity, quality (measured by peer evaluations), and self-evaluation.[1] He found that in 53 per cent of the cases self evaluations and peer evaluations of the same publications were reasonably related.[2] This is insufficient agreement to warrant substituting one for the other.

Rationale for the Particular Method Used in this Study

In a field where research is thought to be less institutionalized, the assumptions required for using an evaluation form for measuring research quality are more likely to be met than the assumptions required for using some other measure of quality. Chief among these assumptions is the requirement of reasonable agreement on the standards for evaluating quality, and reasonable consensus in the way the standards are applied. Finally, there must be agreement among the members of a field that the people making the judgments are competent. Because we have reason to believe that research in education is not highly institutionalized,[1] it seems all the more important to make standards for evaluation explicit, so members of the field can indicate whether or not they agree on the standards, and so we can measure the amount of consensus on those standards.

---

[1] Byron T. Shaw, The Use of Quality and Quantity of Publication as Criteria for Evaluating Scientists, Agricultural Research Service, U.S. Department of Agriculture, Miscellaneous Publication No. 1041 (Washington, D.C.: Government Printing Office, 1967).

[2] *Ibid.*, p. 39.

For all of the above reasons, we have decided that an evaluation form is the most appropriate measure of quality in educational research. The issue then becomes: what type of evaluation form? In Chapter IV we will describe the specific evaluation form used in this study and we will present the rationale for decisions made about content, administration, and scoring.

---

[1] The concept of the level of institutionalization of research in education is discussed extensively in Wilder's study "The Reading Experts." We return to the concept of institutionalization in Chapter X.

CHAPTER IV

ADMINISTRATION OF THE RATING FORM AND AGREEMENT AMONG THE JUDGES

The nature of the Rating Form used to measure quality and the way it is administered may affect the amount of consensus among judges. In this chapter we discuss the Rating Form used in this study and how we administered it. In addition to presenting quantitative data on the amount of agreement, we do qualitative analyses of empirical research reports and of judge characteristics in order to understand better what contributes to greater or lesser consensus among judges.

The Rating Form

Nature of the Form

As we have seen in the previous chapter, the one set of standards which Wandt, Gephart, Nasatir and Persell all agreed upon was an overall, or global, measure of a research paper's quality. In addition, for a study such as the present one, an evaluation form should be reliable, i.e., elicit similar evaluations from the different judges using it, and valid in the sense of corresponding to the concept of quality we have in mind. Finally, it should be easy to administer. We will consider the evaluation form used in this study, and present evidence for accepting it as reliable, valid, and easy to administer.

The rating form used in this study (Figure IV-1) requires an overall evaluation on a five-point scale of each of three dimensions of research quality: (1) substantive contribution to theoretical knowledge in education or a discipline; (2) substantive contribution to any field(s) of educational practice; and (3) utilization of (or contribution to) research methods.

Reliability

A rating form very similar to the one being used in this study was developed by Nasatir[1] (Figure IV-2). The only difference was that the contribution to substance was not divided into theory and practice. Nasatir had ten judges rate six articles on education in common. Using Nasatir's Global Rating Form, Persell had nineteen judges rate four articles in common.[2]

The consensus (as measured by the percentage of judges in the mode) in both these administrations was high (Table VI-1). This was true even though the judges were drawn from different backgrounds and settings. In each case, half were from Teachers College of Columbia University and half were from the Bureau of Applied Social Research, which is affiliated with the Sociology Department of Columbia University. Although these judges were from only two different institutions, we expected similar consensus when the judges came from different institutions around the country. In the next section of this chapter we will examine in detail the evidence bearing on this expectation.

---

[1] Nasatir, "Project Memorandum No. 5."

[2] [Persell], "Measuring Educational Research Quality."

1-2/05 RATING FORM C. Persell 4/69

3-6/ 6-8/

Article Number:____________________ Rater's Number:____________

On the basis of an over-all impression of the research article you have just read, please rate it (using the scale shown below on the left side of the page) as to both substance and methodology, checking one box in each column.

	(A) SUBSTANCE		(B) METHODOLOGY
	The ideas or empirical findings presented; their substantive contribution to any field(s) of theoretical knowledge in education or a discipline 9/	The ideas or empirical findings presented; their substantive contribution to any field(s) of educational practice 10/	The study's utilization of (or contribution to) research methods 11/
(I) Ranks with the best empirical research studies known to me; on a par with the top 5 or 10 per cent in this respect.	(1)	(1)	(1)
(II) Better than average, though not "outstanding" in this respect.	(2)	(2)	(2)
(III) Run of the mill in this respect; neither better nor poorer than the bulk of research that I have seen.	(3)	(3)	(3)
(IV) Not up to "average" standards; "less-than-mediocre" in this respect (although not altogether lacking).	(4)	(4)	(4)
(V) Incompetent in this respect; among the poorest examples of "research" that I have encountered.	(5)	(5)	(5)

FIGURE IV-1

FIGURE IV-2

Rater's Name: ______________________________

Article No.: ______________________________

Date: ______________________

Global Rating Form

On the basis of a quick first impression, a "snap judgment" about the research report you have just read, please rate this study (using the scale shown below on the left side of the page) as to both:

	(a) The ideas or empirical findings presented; their substantive contribution to any field(s) of knowledge or practice:	(b) The study's contribution to (or utilization of) research method:
(I) Ranks with the best empirical research studies known to me; on a par with the top 5 or 10 per cent in this respect.		
(II) Better than average, though not "outstanding" in this respect.		
(III) Run of the mill in this respect; neither better nor poorer than the bulk of research that I've known.		
(IV) Not up to "average" standards; "less-than mediocre" in this respect (although not altogether lacking).		
(V) Incompetent in this respect; among the poorest examples of "research" that I've encountered.		

TABLE IV-1

A COMPARISON OF TWO DIFFERENT ADMINISTRATIONS OF AN EARLIER GLOBAL RATING FORM

		Percentage of Judges in the Mode
Nasatir administration:	Substantive	70%
	Methodological	68%
Persell administration:	Substantive	70%
	Methodological	72%

Validity

The question of validity is the question of whether the operational definition of a concept is measuring what it purports to measure, in this case research quality. We propose two ways of assessing the validity of this measure of quality:

(1) By comparing the nominal and operational definitions of the concept.[1] This method is a logical operation, and we tried to do it in some detail in Chapter III.

(2) By comparing two operational measures of a concept to see if they classify the phenomena being studied in a similar fashion. In the Persell study mentioned above, a forty-six item detailed rating instrument (Appendix IV-1) was used by the judges in addition to the Nasatir Global Rating Form. Thus, two separate operational measures

---

[1] Suggested by Hans L. Zetterberg, On Theory and Verification in Sociology (Stockholm: Almqvist and Wiksell; New York: Tressler Press, 2nd ed., 1957), Chapter III.

were used by nineteen judges on 102 articles from 38 journals. When the articles were ranked according to each of the two measures and those rankings were compared, the coefficient of rank correlation was .81. This result gives us confidence that the relative quality of the articles was deemed similar by each method of measuring.[1]

Ease of administration

The final reason we had for accepting the Rating Form was its ease of administration. In comparison with the Detailed Criteria Rating Form which judges found time-consuming and sometimes difficult to apply in the Persell study,[2] the one-page Rating Form is short and simple. Since there is less agreement among the judges on the Detailed Criteria Rating Form than on the Global Rating Form, ease of administration is not gained at the expense of consensus. Finally, the Rating Form is more easily applied to all designs of research, whether experiments, analytic surveys, or some other type. Certain requirements must be met, however, in the administration of the rating form, and we turn now to a discussion of these.

Method of Administration in This Study

Selection of judges

In this study, each of thirty-nine judges rated eleven research papers on education. A common feature of the earlier work by Wandt,

---

[1] As we suggested in Chapter II, it would be interesting to compare this instrument with other methods of measuring research quality, such as citation rates and peer evaluations.

[2] Persell, "Measuring Educational Research Quality."

Gephart, Nasatir, and Persell[1] was the use of "qualified" judges for rating the research papers. By qualified we mean people who have done research on education that is well regarded by their peers. We sought as judges people who were acknowledged experts in their field. Some, although by no means all, of the judges were field readers of research proposals for the U.S. Office of Education.[2] Others were leaders and officers in the American Educational Research Association (AERA). Of course, all qualified individuals were not asked to be judges. Judges were selected from a range of substantive specialties corresponding to the distribution of paper topics in the population. When someone we contacted was unable to do the ratings, we asked them to recommend someone they respected in the same substantive specialty.[3]

Each judge rated eleven research papers.[4] Wherever possible, judges were assigned papers to rate in one of their substantive specialties. Wandt recommends this practice.[1] Gephart's findings also support this procedure, since he found higher interrater reliability when the jurors had backgrounds "either in the substantive area of the research articles or in its research design and methodology."[2] The judges were told very little about the study in the letter requesting their participation (see Appendix IV-2) or in the directions (Appendix IV-3) accompanying the papers and the Rating Forms. All identifying characteristics of the research papers were omitted on the Xeroxed copies of the research papers sent to the judges. This included the name(s) and location of the author(s), the name of the journal, any indication of funds which might have supported the project on which the paper was based, and any indication that the paper was part of a student's school work. Professional acknowledgments to colleagues which might have indicated where the author was located were also eliminated. Only the title of the paper and its content remained. No rater was assigned papers to rate which were written by people located in his own university.

---

[1] Wandt, et al., Evaluation of Educational Research; Gephart, Development of an Instrument; Nasatir, "Project Memorandum No. 5"; Persell, "Measuring Educational Research Quality."

[2] Field Reader Catalogue, Office of Education, U.S. Department of Health, Education, and Welfare, prepared by Autonetics Division of North American Rockwell Corp., Contract No. OEC-4-6-001305-1017, August 1968.

[3] The judges were: C. Arnold Anderson, Kurt Back, Arno A. Bellack, Carl Bereiter, Charles E. Bidwell, Richard H. Bloomer, Walter Borg, Wilbur B. Brookover, Richard O. Carlson, Jeanne S. Chall, Lee J. Cronbach, Robert Dreeben, Nicholas Fattu, Joshua A. Fishman, Robert Gagne, Gene V. Glass, Miriam Goldberg, David Goslin, David Gottlieb, Robert J. Havighurst, Robert Herriott, Philip W. Jackson, Fred N. Kerlinger, Russell Kropp, C. Mauritz Lindvall, Walter MacGinitie, Donald L. Meyer, Andrew Porter, Lauren B. Resnick, Dale P. Scannell, Pauline Sears, Saul B. Sells, Edward A. Suchman, Bert E. Swanson, David V. Tiedeman, E. Paul Torrance, Ralph W. Tyler, Richard Wolf, Albert H. Yee.

[4] How the population of research papers was defined and selected was described in Chapter III.

We added a Specification Form (Figure IV-3) to the Rating Form to obtain information about whether prior knowledge of a paper affects the judge's rating of it, and to ascertain if judges are more likely to be familiar with the best research.[3] Also the Specification Form

---

[1] Wandt, et al., Evaluation of Educational Research.

[2] Gephart, Development of an Instrument, p. 130.

[3] This idea was suggested by a conversation with Stephen Cole of the State University of New York at Stony Brook.

FIGURE IV-3

C. Persell 4/69 **SPECIFICATION FORM** Article Number________

1. Do you remember having read this article before? 1 ☐ yes 2 ☐ no 12/

2. Do you remember:

   (a) Who the author is? 1 ☐ yes 2 ☐ no 13/

   (b) Where the paper appeared? 1 ☐ yes 2 ☐ no 14/

3. If the article is rated poor (4 or 5) in any of the 3 categories, please answer these two questions.

   A) IF THE ARTICLE IS RATED POOR (4 or 5) ON EITHER SUBSTANCE CATEGORY, is it because the article does not deal with an "important" theoretical or practical problem in the field, or because while it treats an important problem, it adds little to our understanding of the problem?

Check one only { 1_____does not deal with an important problem 15/

2_____although topic represents an important problem, the article adds little to our understanding of it.

3_____other (please specify)________________________________ 16-20/

____

   B) IF THE ARTICLE IS RATED POOR (4 or 5) ON METHODOLOGY, is it because of poor design, method of data collection, or data analysis?

   (Please check all that apply): 21/

1_____poor study design

2_____poor method(s) of data collection

3_____poor method(s) of data analysis

4_____other (please specify)________________________________ 22-23/

____

provides additional information about what detracts from research papers, as we shall see in Chapter V.

Selection of papers rated by several judges

Even though the rating form being used here is similar to Nasatir's Global Rating Form, we wanted to understand its reliability under the circumstances of this study. To do this, we had each of ten research papers on education rated in common by five judges.[1]

The ten papers rated in common were selected in the following manner: The papers were grouped according to primary substance. There were ten major categories of primary substance, namely, administration, curriculum, personnel and teacher training, counseling and guidance, research methods, psychological processes, higher education, reading, social context of education, and speech and special education. Within each of these ten, one paper was randomly selected to be rated in common. This method of stratified random sampling was employed to insure that each major substantive area was represented. After the ten papers were selected, they were scanned, to see if they represented a variety of research methods, research quality, length, and type of journal. Three of the ten papers were given at the 1968 annual AERA meeting. Since it did illustrate variety in these respects, the sample was considered a useful one for the purpose of assessing inter-rater reliability, and was therefore retained.

---

[1] The net result of this operation was the following:
6 papers were rated in common by 5 judges (as planned);
1 paper was rated in common by 6 judges (we gave one judge the wrong paper);
2 papers were rated in common by 4 judges (one judge refused to rate a paper so removed from his specialty, and 1 judge had not done the ratings at the time of this analysis);
1 paper was rated in common by 2 judges (one of these papers was to go to the judge who received the one rated by six people above; in the other two cases, we inadvertently assigned two judges an eleventh paper which was not in the group to be rated in common).
The analysis of consensus in this study is based on ratings of the nine papers for which we had ratings by four or more judges.

Overall Agreement among the Judges

Quantitative Results

Because the number of judges was smaller in this administration, we decided that the percentage of judges in the mode was not the best measure of consensus. A mathematical statistician[1] suggested that the average deviation from the mean rating would be an appropriate way to describe the amount of agreement among judges.[2] If we examine the overall average deviation (obtained by averaging the average deviations obtained on each of the three dimensions) of the ratings,[3] we see that there is generally high agreement among the judges (Table IV-2). In no paper does the overall average deviation exceed 1.0, which would be one adjacent category on the Rating Form away from the mean. The lowest

---

[1] The thoughtful help with this problem that was provided by Robert Berk from the Columbia University Department of Mathematical Statistics is gratefully acknowledged.

[2] Because all nine papers were not rated in common by the nine judges, we were unable to use such an analysis of variance in ranks as Kendall's W, the coefficient of concordance.

[3] A rough rule of thumb for interpreting the average deviation on a five point scale is the following: an average deviation of 0 means complete agreement among all judges, an average deviation of 1.2 would occur if the judges were distributing their ratings across the scale in a completely random fashion, and an average deviation of 2 would indicate complete disagreement among judges.

TABLE IV-2

AVERAGE DEVIATIONS FROM THE MEAN RATING OF NINE RESEARCH PAPERS RATED IN COMMON

	Average Deviation from the Mean Rating on			Overall Average Deviation
	Contribution to Theory	Contribution to Practice	Use of Research Methods	
Paper 876 (4 judges)	.50	.38	.38	.42
Paper 124 (5 judges)	.48	.64	.32	.48
Paper 886 (5 judges)	0	.60	.88	.49
Paper 321 (4 judges)	.38	.75	.75	.63
Paper 1180 (5 judges)	.72	.48	.72	.64
Paper 287 (6 judges)	.48	.80	.80	.69
Paper 527 (5 judges)	.64	.72	.72	.69
Paper 545 (5 judges)	.64	.64	.88	.72
Paper 1018 (5 judges)	.88	1.04	.64	.85
Overall average deviation on each of the three dimensions	.52	.67	.68	.62

overall average deviation is .42 which is less than half an adjacent category from the mean. If we average the overall average deviations across all nine papers, the result is a mean overall average deviation of .62. These results give us confidence in the amount of agreement among judges, and therefore make it reasonable for us to assume that since the judges generally agree, we are justified in using a single rating of a research paper as a measure of its quality. While these findings give us confidence in the reliability of the Rating Form, however, by themselves they tell us little about what contributes to greater or lesser agreement among judges. In order to get some insight into the nature of agreement among judges, we qualitatively analyze papers eliciting the greatest and least agreement among raters.

Qualitative Analysis of the Paper with Greatest Overall Agreement

Paper characteristics

Paper 876 shows the lowest overall average deviation (and thus the highest agreement) among judges, with an average deviation of .42. This paper describes a state department of public health's state-wide Hearing Conservation Program. It presents audiological data on 1,000 children with hearing loss, and it analyzes hearing loss as a function of age, type of loss, and audiometric configuration. Also, it makes observations relating hearing loss and noise exposure, regression of sensori-neural loss, and Hz sensori-neural "spikes." Methodologically it is primarily descriptive, presenting percentage differences in various nominal categories. The paper does not indicate how the 1,000 children described were selected, although it does say that they came from ten counties representing mountain and plain regions, and low, medium, and high income per capita groups. In general terms, hearing loss represents an important practical problem with implications for education. The paper relates its study to relevant literature, but it does not refer to theory in a way that increases our understanding of the problem. In Table IV-3 we can see how each judge rated the paper on each of the three dimensions of quality. The two judges rating the paper "below average" (4) in contribution to theory both specified their rating by checking, "although the topic represents an important problem, the article adds little to our understanding of it." We agree with this evaluation. The one judge rating the paper below average (4) in research methods checked "poor study design" and "poor methods of data analysis" as the reasons for the rating.

TABLE IV-3

JUDGES' RATINGS OF PAPER 876, THE ONE WITH HIGHEST OVERALL AGREEMENT AMONG RATERS

	Judges' Rating of		
	Contribution to Theory	Contribution to Practice	Use of Research Methods
Judge 03	3	3	3
Judge 16	4	3	4
Judge 17	4	2	3
Judge 21	3	3	3

Altogether, both the judges' ratings and their comments indicate their evaluation of this paper as mediocre. It is a bit surprising to me that the paper with the highest agreement is a mediocre article, in view of Nasatir's and Persell's[1] findings that generally the papers at the extremes of quality, i.e., either very good or very poor, elicit the highest consensus. This raises the question, is quality related to agreement? Clearly it is not in the case of paper 876, but perhaps it is for other papers (Table IV-4). The papers which rank second and third in agreement among judges are at the extremes of quality so the tendency for quality to affect agreement is true in this administration of the rating instrument as well as in the earlier Nasatir and Persell studies. That paper 876, which elicits highest agreement, is not at the extremes of quality suggests that there are other factors which may also be related to agreement.

One of the characteristics that may affect agreement is similarity between the judges' substantive specialty and that of the paper. This was found in the administrations of other rating instruments[2] to contribute to consensus among judges. In the case of paper 876, however, this was not the case either. The paper was not substantively similar to the specialties of any of the judges. The fifth judge who was supposed to rate this paper refused to rate the paper altogether, although asked twice, on the grounds that it was not in

---

[1] Nasatir, "Project Memorandum No. 5"; Persell, "Measuring Educational Research Quality."

[2] Nasatir, "Project Memorandum No. 5"; Gephart, Development of an Instrument; Persell, "Measuring Educational Research Quality."

TABLE IV-4

QUALITY AND AGREEMENT AMONG JUDGES

Papers Ranked by Research Quality (Mean Quality Rating on Each Dimension)					Overall Average Deviations from the Means on 3 Dimensions	Rank on Agreement among Judges
		Theory	Practice	Methods		
ABOVE AVERAGE	Paper 886	2.0	2.2	3.2	.49	3
AVERAGE	Paper 545	3.6	2.8	2.4	.72	8
	Paper 287	2.4	3.0	3.0	.69	6.5
	Paper 876	3.5	2.8	3.3	.42	1
	Paper 1018	3.6	3.2	2.8	.85	9
	Paper 527	3.2	3.4	3.4	.69	6.5
	Paper 321	3.8	3.3	3.3	.63	4
BELOW AVERAGE	Paper 1180	3.6	3.6	3.6	.64	5
	Paper 124	4.6	4.2	4.2	.48	2

her area of expertise. Another judge refused initially to rate it, and only did so reluctantly, with the comment, "I don't consider myself competent in this area of research." This suggests that substantive similarity between the paper and the judges' specialties was not related to the high agreement obtained on this paper. In fact it is difficult to find any characteristic of the paper which seems to explain why it evinced such high agreement among judges.

Judge characteristics

If there is not something apparent about the paper which accounts for high agreement, perhaps there is something about the judges which contributed to their consensus. Are they, for example, a particularly homogeneous group? As we just mentioned, they are not homogeneous with respect to their specialties, which range over special education, evaluation, curriculum, and sociology of education. Nor are they particularly alike with regard to their organizational affiliation. One is in a liberal arts department, one is in a research institute, and two are in schools of education. It seems that they are not similar in terms of specialty or setting. Regarding their training, two were completely trained in the liberal arts, and two were trained in both the liberal arts and education, so they do not seem similar in training either. Their ages range from 50 to 69 and they are all males, but it seems doubtful that either of these factors accounts for their agreement. The only characteristic they seem to share is lack of familiarity with the paper. Three of the judges do not remember having read the paper before, who the author is, or where the paper appeared, and one did not answer those questions. But it seems unlikely that familiarity affects consensus.

It is conceivable that the high level of agreement is related to the fact that only four judges ended up rating this paper. If we look at Table IV-2 again, however, we see that another paper (321) which was rated by only four judges did not have as high consensus (the overall average deviation for that paper was .63) and that one paper (287) which was rated by six judges elicited average agreement among judges (the overall average deviation was .69), so the amount of agreement does not seem to be affected by the number of judges rating the paper. In short, none of the paper or judge characteristics which we considered seem to be related to the high agreement among judges.

Qualitative Analysis of the Paper with Least Overall Agreement

Paper characteristics

The paper which yielded lowest agreement among the judges was 1018. This paper considers which method of laboratory instruction is more effective in attaining the objectives of a biology course for college freshmen--a "content-centered" lab with student activities revolving around oral and written directions or a "process-centered" laboratory with students applying "inductive methods" to the solution of major problems which have been identified by either the students or the teacher. The study was designed as follows: of twenty freshman college biology sections in one college, three were arbitrarily designated "content-centered" and three "process-centered." Each group of three was given the treatment typical of the particular method. Each group was also given a battery of before and after tests to measure the objectives of increasing knowledge of biology, developing scientific attitudes, developing problem-solving abilities, and developing an interest in biology. Analysis of variance, analysis of co-variance, and correlations were used to measure differences in the two groups. We think the problem is of practical significance for higher education. Also it is our feeling that this is a better paper than 876 (the one eliciting such high consensus), but the overall ratings of the judges do not reflect this (Table IV-5).

TABLE IV-5

JUDGES' RATINGS OF PAPER 1018, THE ONE WITH LEAST OVERALL AGREEMENT AMONG RATERS

	Judges' Rating of		
	Contribution to Theory	Contribution to Practice	Use of Research Methods
Judge 10	3	2	3
Judge 20	2	2	2
Judge 11	4	3	2
Judge 09	4	4	3
Judge 28	5	5	4

Three of the judges rated this paper average or above, and two rated it below average. The judges rating it below average made a number of comments which illustrate their reactions to the paper. The most general comment about the paper (and the only comment made by a judge rating the paper average or above) was made by one judge who wrote, "If I had made a decision at or about page 3, I would have rated contribution to theory and contribution to practice as above average. Later pages detracted from this impression." The early pages of the paper state the problem and the hypotheses to be tested. We might expect that there would then be some connection with theory which might suggest the anticipated direction of findings. But as one judge observes, the paper "does not deal with [a problem] in any theoretically important way." Similarly, another judge indicates, "Although topic represents an important problem, the article adds little to our understanding of it." A third judge is in agreement when he writes, "does not deal with a theoretical problem." He also writes that there is "no such [theoretical] knowledge." We wish he had elaborated on this comment, because it is hard to tell exactly what he means. Is there no theory, or even general orientation, which suggests that one type of curriculum would lead to more learning than another type of curriculum? If not, this would seem to be a problem a research study in the area would try to remedy.

The more severe judges also comment about what is wrong with the methodology of this paper. Two judges mention poor methods of data collection and data analysis. One of them writes, "Statistics OK--need item analysis." Also, "Poor reporting of homemade instrument." We agree with this comment. The paper spends more time reporting the reliability and validity of published instruments than of instruments especially designed for the study. The other judge also checks poor study design, and writes, "He is unable to set up an experiment really testing the basic hypotheses."

Judge characteristics

The differences in these judges' ratings is puzzling because in terms of their specialties, educational backgrounds, and locations these judges are more homogeneous than the ones rating paper 876, which elicited the greatest agreement. Four out of five of these judges specialize in curriculum, and one is in research methods. While the judges are not specialists in science education, they are closer to the substance of the paper than in the case of paper 876. Four out of five received two or more of their degrees in the liberal arts. Only two had degrees in education. Three out of five are in (or retired from) research institutes, with the other two in schools of education. All are male and their ages range from 42-65. None of the judges remember having read the paper before, who the author is, or where the paper appeared. In brief, it seems that none of the background characteristics we know about these judges explain any of the variation in agreement on this paper, just as they did not explain the high agreement on paper 876. What seems to be the case here is that there are two judges who are very severe in their rating of this research paper. Are these judges consistently severe in their ratings of research papers, or is their severity limited to this particular paper? We can try to answer this question, even without having other ratings in common done by these same judges, by examining the distribution of each judge's ratings, to see if some judges are consistently more severe or more lenient in their ratings. The ratings of eight judges (starred in Appendix IV-4) seem more severe than those of the thirty-one other judges, and three seem more lenient (marked with a "t" in Appendix IV-4). Variation in the distribution of judges' ratings does not concern us, however, unless it is accompanied by great disagreement among judges. It may reflect the actual distribution of quality of the papers rated by that judge. If a judge is very different from his colleagues, however, we cannot help wondering if there is something peculiar to that judge which makes his ratings more severe.

The two judges who rated paper 1018 more critically than the other three judges tended to have ratings that were more severe than those of other judges, particularly with regard to theory for one judge and theory and methods for the other judges. Since it is impossible to tell if other judges would have rated those papers as harshly, however, we decided not to eliminate the ratings of these two judges. Further, if we had excluded their ratings of twenty papers, we would have had too few cases for most of our three-variable analysis.

Agreement among the Judges on Different Dimensions

Agreement on Contribution to Theory

Table IV-2 reveals that the dimension of research quality that generally elicits most agreement among the judges is contribution to theory. This is particularly interesting in view of the diversity of the judges and of the substantive areas of the papers. How can we account for this relatively high agreement on contribution to theory ratings? First, theory may be present or absent in this sample of papers. In contrast to theory, some type of research methods was required for inclusion in this sample. Thus, it is quite possible that theoretical treatment of a problem may be completely absent from a paper. If that is the case, it would be relatively easy for judges to agree that a paper was below average with regard to theory. The relatively greater agreement among judges on ratings of theory may be due to a shared view among judges about what theory is, and about how a paper contributes to theory. We will keep these possibilities in mind as we examine the papers showing the most and least agreement on ratings of contribution to theory.

Qualitative analysis of papers and judges with most and least agreement on contribution to theory

Paper 886 shows the highest agreement among all the ratings of theory. The problem that this paper treats is the role that education, and specifically educational selection, plays in the formation of the social and political attitudes of those who are educated. The paper presents three theoretical orientations about ways in which an educational system may affect social and political attitudes. First, the paper suggests that education may limit choice through its close relationship to the social stratification system. This statement is then related to theories of social stratification. Second, an educational system may influence attitude formation through the pattern of authority within schools. This idea is related to theories of political organization. Finally, educational systems are used to transmit social and political values. Five judges all rate this paper above average (2) on contribution to theory. The high agreement among judges may well be due to the fact that the paper is better than average, and therefore toward one extreme of the spectrum of quality.

There may also be characteristics of the judges, however, which contribute to their agreement on this paper's contribution to theory. All of the judges were trained in the liberal arts, all but one of them are specialists in the sociology of education, and that one is a specialist in research methods. Two of them are now sociologists in liberal arts departments, another two are sociologists in a department of education, and the fifth is at a foundation. So we see that these judges are quite homogeneous both in relation to the substance of the paper and in relation to each other. We have seen earlier that homogeneity of the judges did not seem to be related to agreement, so how can we explain its apparent relationship here? The distinguishing feature in this case seems to be that all of these judges are not only in the same specialty, but their specialty is a behavioral science discipline. We expect to find in this study that research is more institutionalized in the behavioral sciences than in education. In view of this expectation, we would also anticipate finding more consensus on research standards among judges in the behavioral sciences than among those in education, and this paper bears out our expectation.

The paper illustrating least agreement among the judges in their evaluation of contribution to theory is 1018. We have already seen the distribution of the judges' ratings on this paper (Table IV-5). The three judges rating the paper below average had similar comments to make about it. One writes, "The paper does not deal with the problem in any theoretically important way." Another writes that the paper "does not deal with a theoretical problem," and the third suggests that theory is not brought to bear on the problem. For the judge rating this paper above average (2), the testing of hypotheses may constitute sufficient contribution to theory to warrant an above average rating, while the other judges require more than that. The low level of agreement on this paper may be due to differences in what judges accept as theory.

As we saw earlier, the judges rating this paper are also quite similar in terms of their specialties, educational backgrounds, and locations. However, they are similar in sharing the specialty of curriculum, an educational specialty rather than a behavioral science specialty. So it seems that homogeneity with regard to a behavioral science specialty is related to agreement among judges, while homogeneity in an educational specialty is not necessarily related to agreement.

Agreement on Contribution to Practice

The papers with the most and least agreement on contribution to practice are the same as the ones with most and least overall agreement (papers 876 and 1018 respectively).

A factor which may be related to agreement among judges rating contribution to practice is how close they are to educational practitioners. Degree of closeness to practitioners might be measured several ways but we have data to measure it only by comparing judges who are in a school of education with those who are not, which may be used as an indicator of closeness to practitioners. We will keep this attribute in mind as we consider the papers showing most and least agreement on contribution to practice.

Qualitative analysis of papers and judges with most and least agreement on contribution to practice

Perhaps paper 876 shows high agreement on contribution to practice because an average contribution to practice is fairly easy to agree upon. Paper 1018, on the other hand, may show low agreement on the rating of contribution to practice because judges do not agree that it contributes to practice (Table IV-5). Some may be confusing their impression that the topic represents an important practical problem with their assessment of how much it contributes to the solution of that practical problem. This might account for the low agreement. As we have already seen, two out of four judges who agree most are in schools of education and two out of five who agree least are in schools of education. From this we conclude that being in a school of education is either not a very good indicator of closeness to practitioners, or that it does not account for the variation in agreement among judges on contribution to practice. The only judge characteristic which we could find associated with agreement was severity on the part of several judges, which seems related to the low agreement on paper 1018.

Agreement on Use of Research Methods

Clearly there are different traditions of research methodology, both in design and analysis, ranging over tests and measurement, experimental designs, analytic surveys, factor analysis, item analysis, analysis of variance, multivariate tabular analysis, path analysis, and others. Rating different methodological traditions with agreement requires shared technical expertise, which is somewhat more specific than shared definitions of theory. Yet it is rare that individual researchers are competent in all types of research methods. For papers in the rating sample, we tried where possible to assign them to judges in the same methodological tradition. For the papers rated in common, however, it was much less possible to assign papers to judges in the same substantive and methodological specialty. Therefore, the relatively lesser agreement among judges on this dimension may be due to heterogeneity in methodological traditions. As we do a brief qualitative analysis of these papers, we can consider this possibility.

Qualitative analysis of papers and judges with most and least agreement on use of research methods

Paper 124 has the most agreement among judges on use of research methods. This paper compares disciplinary problem students and a random sample of other college dormitory residents on thirty factors. There were fifty-nine disciplinary students and fifty-nine others randomly selected from the remaining residents of a dorm at a large state university. Chi-Square was used to compare the two groups on data for thirty intellective and non-intellective factors. The judges are in strong agreement on their rating of methodology, and also in the reasons they cite for its being regarded as below average (Table IV-6). Clearly there is something about this paper that catches the attention of the judges in specifying their below average rating of the paper on methods. The design is hardly an elaborate one, and the method of analysis, Chi-Square, is learned early by students in educational fields and in sociology. Since the judges' specialties are educational psychology, counseling and guidance, educational administration, and sociology of education, it is reasonable to assume that they are all familiar with the type of design and analysis used in this paper, and thus able to agree on whether or not it is an appropriate method and one that was executed soundly. Had the design or method of analysis been less common, they might have been less likely to agree.

TABLE IV-6

JUDGES' RATINGS OF PAPER 124, THE ONE WITH HIGHEST AGREEMENT ON RESEARCH METHODS

	Rating of Use of Research Methods	Comments
Judge 31	4	Poor methods of data analysis
Judge 22	4	Poor design, poor methods of data analysis
Judge 04	4	Poor design, poor methods of data analysis
Judge 07	4	Poor design
Judge 01	5	Other (illegible comment)

The judges rating this paper are also quite similar in training and present location. Three out of five were trained in the liberal arts and the other two were trained in both the liberal arts and education. Four out of five are in schools or departments of education and the fifth is in an educational research institute.

Paper 886 evinces the least amount of agreement among the judges on use of research methods. It examines the relationship of educational selection to social and political attitudes. The data were collected by interviews with 331 14- and 15-year-old English schoolboys in one public, three grammar, three secondary modern, and three comprehensive schools. In addition, it is based on secondary analysis of a Gallup Youth Survey, a panel study of a random area sample of 1,800 British boys and girls from 16 to 18, done in four waves in 1960-63. The judges' ratings of the paper's use of research methods are quite consistent with one major exception (Table IV-7). Since the other four judges generally agree, it is impossible not to ask what is it about the one judge which makes his rating so different? He comments that there are too few cases and that simplistic logic is used. We agree[1] that in one table there are too few people in some of the categories (7 out of 24 categories in that table had 20 people or fewer in them), but there are ten other tables in the paper, all of which have considerably higher N's. We think it is too harsh to rate the paper incompetent on methodology for this reason. This judge seems to be another example of a severe rater. There is nothing in what we know of his background to account for this. The major indicator we have of his severity are his many, often illegible, and a bit unusual comments scrawled on all of his rating sheets. For example, along the bottom of the Rating Form for this paper was written "and class obsessed [or oppressed]." (Because of this judge's scribbled and harsh comments we asked some of his students and colleagues what he was like. They informed us that he is extremely opinionated and critical in his views of research studies.) The other four judges are very homogeneous, both in their background and in relation to the paper. They received all of their degrees in the liberal arts, three out of four are now in liberal arts departments, and the fourth is with a foundation. Three are specialists in sociology and one in research methods.

---

[1] In an effort to better understand the rating process, we indicate where we agree and disagree with the judges' evaluations throughout the qualitative analyses presented here. Of course, our opinions never alter the judges' ratings of research quality analyzed later in this volume.

TABLE IV-7

JUDGES' RATINGS OF PAPER 886, THE ONE WITH LEAST AGREEMENT ON RESEARCH METHODS

	Rating of Use of Research Methods
Judge 26	2
Judge 24	3
Judge 15	3
Judge 02	3
Judge 01	5

Agreement on the Mean Score

The Mean Quality Score was formed by averaging the scores on each of the three dimensions of quality. Thus, if a paper was rated average (3) on contribution to theory, average (3) on contribution to practice, and better than average (2) on use of research methods, it would have a Mean Quality Score of 2.7.

We can measure agreement on the Mean Quality Score by computing the average of the deviations from it. This measure differs from the overall average deviation which we considered earlier. The overall average deviation is the average of the deviations on each of the three dimensions of research quality, which was presented in Table IV-2. If we compare the average of the deviations from the Mean Quality Score and the overall average deviation across the three dimensions, we see that there is less deviation from the Mean Quality Score (Table IV-9) for each paper and altogether.

TABLE IV-9

COMPARISON OF DEVIATIONS ON THREE DIMENSIONS AND ON MEAN SCORE

	Average Deviations Over All 3 Dimensions	Average of the Deviations from Mean Score
Paper 224	.48	.40
Paper 287	.69	.56
Paper 321	.63	.47
Paper 527	.69	.67
Paper 545	.72	.51
Paper 876	.42	.30
Paper 886	.49	.36
Paper 1018	.85	.77
Paper 1180	.64	.50
Average over all papers	.62	.50

One reason this may be the case is that the effect of one deviant rating, as in 01's rating of research methods on paper 886, is weakened somewhat (Table IV-9).

TABLE IV-9

JUDGES' RATINGS OF PAPER 886

Ratings on:			Mean Score			
Theory	Practice	Research Methods				
2	1	5		2.7		
2	2	2		2.0		
2	2	3		2.3		
2	2	3		2.3		
2	4	3		3.0		
			Average Mean Score	2.5	Average Deviation from Mean Score	.36

While one judge rated 886 incompetent (5) on use of research methods, he rated it higher on practice than anyone else did so his deviation from the ratings of other judges was somewhat reduced in the Mean Quality Score. Higher agreement on the Mean Quality Score undoubtedly also reflects the phenomenon of index formation.[1] There is a higher probability of consensus on an index of several interrelated items than on each of the individual items alone.

---

[1] See Paul F. Lazarsfeld, "Evidence and Inference in Social Research," Daedalus, LXXXVII, No. 4 (1958); also Bobbs-Merrill reprint S-441.

Summary

The administration of the one page Rating Form in this study shows considerable consensus among judges rating research papers in common. The overall average deviation on all three dimensions and across all groupings of judges was .62, which is less than two-thirds of a category away from the mean score. The smallest deviation was 0 and the largest deviation exceeded .88 only once. This evidence gives confidence in the reliability of the Rating Form. Furthermore, the Rating Form is easy to administer.

When we examined paper characteristics which might be related to agreement on the overall rating, we learned that agreement tends to be somewhat higher on papers at the extremes of quality. Examining judge characteristics which might be related to agreement, we could find only one, severity on the part of individual judges, which seems to be negatively related to agreement.

Turning to agreement on the individual dimensions of research quality, we found several interesting relationships. Agreement with regard to theory seems more likely in papers which have relevant bodies of theory clearly specified and related to the problem of the study. Conversely, the paper which presented theory ambiguously and which contained hypotheses to be tested without relating them to a larger body of theory, showed the least agreement among judges on contribution to theory. The paper with the highest agreement was rated by judges in the same behavioral science specialty. Conversely, the paper with the least agreement was rated by judges who were not specialists in the same behavioral science field. It seems that homogeneity with regard to a behavioral science specialty is related to agreement among judges, while homogeneity in an educational specialty is not necessarily related to agreement.

Regarding consensus on contribution to practice, no paper characteristics seem related to agreement, and the only judge characteristic which seems related to agreement is severity. On use of research methods, consensus may be related to use of methods familiar to all of the judges. Low agreement is apparently unrelated to paper characteristics. Severity of particular judges is also negatively related to agreement on ratings of use of research methods.

Finally, in this chapter we saw how agreement on the Mean Quality Score was even higher than agreement on individual dimensions. Since agreement is high on both measures, we will generally examine quality on each of the three dimensions separately.

CHAPTER V

THE DISTRIBUTION OF RESEARCH QUALITY IN THE POPULATION AND THE RELATION OF THE DIMENSIONS OF RESEARCH QUALITY TO EACH OTHER

Distribution of Research Quality in the Population

In this section we will describe the distribution of research quality on each of the three dimensions in the Rating Form--contribution to theory, contribution to practice, and use of (or contribution to) research methods.

Distribution of Ratings of Contribution to Theory

The largest percentage of papers were rated "incompetent" in terms of theory (Table V-1).

TABLE V-1

DISTRIBUTION OF RATINGS OF CONTRIBUTION TO THEORY

	Rating					
	Best	Above Average	Average	Below Average	Incompetent	
	(1)	(2)	(3)	(4)	(5)	
Percentage of cases	4%	26	27	25	18	= 100%
Number of cases	16	101	105	98	70	= 390

Nearly one in five (18 per cent) of the papers were rated "Incompetent" with regard to their substantive contribution to any field(s) of theoretical knowledge in education or a discipline. An additional 25 per cent of the papers were judged "Not up to average standards; less-than-mediocre" with respect to their contribution to theory. The amount of research rated very good does not off-set this generally unfavorable evaluation of contribution to theory. Only 4 per cent of the papers were considered to "rank with the best empirical research studies known to me; on a par with the top 5 or 10 per cent in this respect."

If a judge rated a paper "Below average" or "Incompetent" on contribution to theory, or contribution to practice, we asked him to specify whether it was "because the article does not deal with an 'important' theoretical or practical problem in the field, or because while it treats an important problem, it adds little to our understanding of the problem." (See Figure IV-3.) It would have been more helpful to ask this question separately for theory and practice to see if the paper was weak with respect to its contribution to theory or practice, but the results are still useful for suggesting where a paper's weakness lies. A much larger percentage of papers are treating an important problem but adding "little to our understanding of it" than failing to deal with an important problem (59 per cent compared to 25 per cent in Table V-2). This finding suggests that the judges feel many researchers are working on important problems, but for some reason are not increasing our understanding of them.

TABLE V-2

JUDGES' EVALUATIONS OF WHY PAPERS ARE POOR ON SUBSTANTIVE CONTRIBUTION

	Reasons			
	Important Problem but Adds Little to Our Understanding of It	Does not Deal with an Important Problem	Other	
Percentage of cases	59%	25	16	= 100%
Number of cases	105	44	29	= 178

If we look at the supplementary comments of the judges about substantive contribution, we find that 60 per cent of the comments concern theory (Table V-3).

TABLE V-3

SUMMARY OF JUDGES' COMMENTS ABOUT THEORY

Substance of Comment	Frequency of Citation
Paper doesn't tie into theory, or treat a theoretical problem	12
Paper adds little to theory, but has practical importance; paper is concerned primarily with a practical problem	7
Paper doesn't produce knowledge of theoretical significance	1
There is an absence of real theoretical knowledge in this area of research (teacher training)	1
	N = 21

The above comments suggest that papers in this sample often fail to connect with theoretical problems. It seems from the above evidence (Table V-2) that many of the papers have the potential for making a contribution to important theoretical problems, but they fail to accomplish this. Throughout this study we will try to understand what facilitates or impedes the production of good research, including that which contributes to theory.

Wandt had 125 judges evaluate a sample of 125 research articles from 41 journals in 1962 on twenty-five characteristics.[1] Two of these characteristics are relevant to the problem of a research paper's contribution to theory. Both reveal significant differences in the ratings assigned articles in education journals and in "related professional journals" (mainly psychology, sociology, and speech) (Table V-4). Both also offer evidence to support one of the hypotheses of this study, namely that better research on education is done in the behavioral sciences than in education.

TABLE V-4

WANDT'S[a] RATINGS RELATED TO THEORY

Characteristic	Mean Ratings (5 is the best rating)		
	All Articles	81 Articles in Education Journals	44 Articles in Prof'l Journals
Problem is significant	3.59	3.31	4.09
Relationship of the problem to previous research is made clear	3.13	2.60	4.02

[a] Wandt, et al., Evaluation of Educational Research, p. 5.

---

[1] Wandt, et al., Evaluation of Educational Research.

Distribution of Ratings of Contribution to Practice

In a practice-oriented field like education, we might assume that even if research is not closely tied to theory it will contribute to educational practice. But this is not the case. Although the distribution of papers according to their contribution to practice is not as bottom heavy as their distribution by contribution to theory, the overall picture is not much brighter than it is for theory; 11 per cent are rated "Incompetent" (Table V-5).[1]

TABLE V-5

DISTRIBUTION OF RATINGS OF CONTRIBUTION TO PRACTICE

	Rating					
	Best (1)	Above Average (2)	Average (3)	Below Average (4)	Incompetent (5)	
Percentage of cases	6%	25	34	24	11	= 100%
Number of cases	22	96	133	95	44	= 390

Every paper does not have to deal with practice or contribute to it. But we are assuming that a good research paper will contribute to either theory or practice. As we will see later in this chapter

---

[1] One of the judges raised the interesting question of whether this group of judges was the most competent for assessing contribution to practice. Practitioners might be better in his opinion. Anyone doing this type of evaluation in the future might consider this possibility. Perhaps the larger percentage of average ratings reflects the judges' uncertainty about judging contribution to practice. None of them reported feeling unable to make such a judgment, however.

(Table V-13), the two dimensions are very strongly related. There are no papers which are rated "Best" on one dimension and "Incompetent" on the other, and very few which are rated "Better than Average" on one and "Below Average" on the other.

Distribution of Ratings of Use of Research Methods

The distribution of papers according to their use of research methods lies between the distribution by theory and by practice (Table V-6). Fourteen per cent were rated "Incompetent" and 5 per cent were rated "Best." For the 39 per cent of papers that were rated "Incompetent" or "Below Average" on use of research methods, we asked the judges to indicate what made the paper weak--poor study design, poor method(s) of data collection, poor method(s) of data analysis, or other (open ended). They could check all that applied.

TABLE V-6

DISTRIBUTION OF RATINGS OF USE OF RESEARCH METHODS

	Rating on Use of Research Methods					
	Best	Above Average	Average	Below Average	Incompetent	
	(1)	(2)	(3)	(4)	(5)	
Percentage of papers	5.1%	24.2	31.1	25.2	14.4	= 100%
Number of papers	20	94	121	98	56	= 389[a]

[a] One judge was unable to rate one paper on its use of research methods, so wherever this variable appears it will have one less case than the other two dimensions of research quality.

The most frequently cited reasons were poor method(s) of data analysis (25 per cent of the papers) and poor study design (24 per cent). Other reasons and poor method(s) of data collection followed (Table V-7).

TABLE V-7

JUDGES' REASONS FOR RATING A PAPER POOR ON RESEARCH METHODS

Reason	% Judges Indicating It
Poor data analysis	25%
Poor study design	24
Other (such as "poor reporting," "confused," "poor interpretation," "analysis did not go far enough")	19
Poor data collection	15
(N = 389 ratings. One paper may have more than one reason indicated.)	

The judges made thirty qualitative comments about what was wrong with the methods of data analysis, the most frequently indicated area of weakness (Table V-8).

The judges made twenty-eight comments about shortcomings in design and methods of data collection. Since these two areas overlap we will consider them together (Table V-9).

These comments of the judges about shortcomings in data analysis and research design suggest that certain fairly specific knowledge about research methods is not being applied by the authors of some research papers. For example, factors are not always controlled, which might account for variations in the dependent variable. Sometimes researchers examine an unrepresentative sample or one that is too small. Whatever the explanation for these shortcomings, they lower the quality of research being done on education.

TABLE V-8

SUMMARY OF JUDGES' COMMENTS ON SHORTCOMINGS IN DATA ANALYSIS

Area of Shortcoming	Frequency of Citation
Poor or inappropriate techniques (e.g., "an obvious example of t-testing repeatedly at great risk of error instead of performing a comprehensive analysis of variance")	11
Poor or inappropriate interpretation (e.g., "tables do not support some of the conclusions")	11
Incomplete analysis (e.g., "the analysis did not go far enough")	8
	30

TABLE V-9

SUMMARY OF JUDGES' COMMENTS ON SHORTCOMINGS IN RESEARCH DESIGN AND DATA COLLECTION

Area of Shortcoming	Frequency of Citation
Inappropriate or defective design (e.g., "too many factors uncontrolled")	14
Inappropriate or too small sample (e.g., "sample was one of convenience and nothing else")	10
Inappropriate or inadequate instruments (e.g., "instrument developed for the study was very weak--carelessly developed")	2
Other	2
	28

We turn now to the other types of weaknesses the judges mentioned (Table V-10). The single biggest reason for a research paper being rated below average with respect to its use of research methods was "poor or sketchy reporting." This weakness was cited thirty-two times by the judges. They repeatedly mention that they cannot tell what the methods were, how the conclusions are related to the findings, or how the findings were generated from the methods. Some researchers may not have learned to specify their procedures in enough detail so they could be replicated. This sketchiness could be due to journal policies, if they do not require procedures to be described fully or if they cut papers so much that they exclude important procedural details. Poor reporting may result from the absence of training in good writing. Graduate schools do little to teach researchers how to write clearly. If a budding researcher does not have the good luck to have an advisor who can write and teach him how to write, he may never have the chance to learn.

TABLE V-10

SUMMARY OF OTHER COMMENTS BY JUDGES
ABOUT RESEARCH METHODS

Nature of Comment	Frequency of Citation
Poor or sketchy reporting	32
General comments on other weakness in a research paper	18
Specific methodological suggestions	8
Emotional comments (e.g., "Awful stuff! Just nothing--and published")	5
Comments about general methodology	3

All of these comments by judges which we have reported are quite similar to those reported by Wandt[1] (Table V-11). He also found that the experts most frequently cited shortcomings in data analysis. Wandt's conclusions about quality in 1962 seem just as applicable today, namely that a large percentage of educational research articles contain serious flaws.[2] These findings reinforce our desire to understand what processes affect the quality of educational research.

TABLE V-11

WANDT'S[a] FINDINGS OF SHORTCOMINGS MOST FREQUENTLY CITED BY EXPERTS IN EVALUATING 125 EDUCATIONAL RESEARCH ARTICLES PUBLISHED IN 1962

Area of Shortcoming	Frequency of Citation
Results of analysis not clearly presented	33
Incorrect methods used to analyze data	32
Inappropriate or defective design	31
Validity and reliability of the evidence not established	29
Conclusions not substantiated by the evidence	29

[a] Wandt, *et al.*, *Evaluation of Educational Research*, p. 6.

Distributions of Ratings on the Three Dimensions Compared

Having examined the distribution of ratings on each of the individual dimensions of research quality, we turn now to a comparison of the three distributions (Table V-12). For clarity in Table V-12, we have condensed the five-point rating scale into three categories.

---

[1] Wandt, *et al.*, *Evaluation of Educational Research*, p. 6.

[2] *Ibid.*, p. 7.

TABLE V-12

A COMPARISON OF THE RATINGS ON THE THREE QUALITY DIMENSIONS

	Rating			N
	Above Average	Average	Below Average	
Contribution to theory	117	105	168	= 390
Contribution to practice	118	133	139	= 390
Use of research methods	114	121	154	= 389

There is very little difference in the distribution of quality along each of the three dimensions (Table V-12). A slightly larger number of papers are rated "Below average" with respect to theory than to either of the other dimensions. There is almost no difference in the number of papers rated "Above average."

This similarity in the three dimensions of quality raises the interesting question of how the three dimensions are related to each other on individual papers. Is there some kind of halo effect operating, or are certain papers good on one dimension and poor on another? In order to answer this question, we will examine the relation between these dimensions separately, and then the relationship among all three simultaneously. Again for clarity we will condense the five-point rating scale into three categories.

Relationship of Theory and Practice Ratings

Quantitative Relationship

The distributions of ratings on each dimension tell us nothing about whether the same papers are rated similarly on theory and practice. There is a high correlation between the theory and practice ratings (Table V-13). Their Tau Beta is .65, which indicates a strong positive relationship. It is significant at the Chi-Square level of .001.

TABLE V-13

RELATIONSHIP BETWEEN THEORY AND PRACTICE RATING

Rating on Contribution to Practice	Rating of Contribution to Theory			N
	Above Average	Average	Below Average	
Above average	84	26	8	[illegible]
Average	28	61	44	133
Below average	5	18	116	139
	117	105	168	= 390

Qualitative Analysis of Papers Rated "Above Average" Theoretically and "Below Average" with Respect to Practice

Five papers were rated "Above average" with respect to theory and "Below average" in reference to practice. These papers warrant qualitative analysis to enhance our understanding of both the rating process and the nature of good research.

Paper characteristics

The judges feel that two of these papers do not deal with an important problem for educational practice. Two others are only inferentially related to educational practice, and one seems very remote from educational practice.

Paper 466 deals with an aspect of learning. The study independently varies the absolute number of associations to be learned and the number of possible pairs of dimensions from which the subject must select the one pair which is relevant to the solution. Theoretically the paper ties in with previous research in the area, and tests an unanswered question in that work. We agree with the judge that it is "Above average" on theory. Specifying his low rating of contribution to practice, the judge indicates, "The paper does not deal with an important problem [for practice]." We agree. Given all the pressing problems facing education, this does not seem to be one of the most important ones.

Paper 615 is a similar example of a study which is related to previous research and theoretical ideas and tries to extend them, but is seemingly unconcerned with the practical significance of the problem being treated. It investigates the influences of negative instances on concept attainment, particularly characteristics of the negative instances themselves. We agree with the judge who specifies his "Below average" rating of contribution to education practice by checking, "Does not deal with an important problem."

Paper 585 also seems quite remote from educational practice. It is concerned with two ways of studying students' judgments about educational objectives--one a multivariate study of students' preference ratings for a set of twenty-four instructional objectives and the other a multivariate study of students' classificational judgments about the content similarity for the same twenty-four objectives. The judge sees this paper as contributing theoretically to statistics. We agree with the rater that the study contributes little to educational practice. The study, by its own admission, is concerned primarily with methodological issues. The judge writes, "The study seems to have been carried out by someone who has learned a lot of statistics but doesn't understand a thing about education and educational objectives. The material in Table 1 [of the paper] reveals an incredible naivete!" We agree. The paper seems unconnected to educational reality.

Papers 932 and 941 are even more removed from educational practice. Paper 932 treats the relation of anxiety to masculinity-femininity in pre-adolescents. It is primarily a study of personality variables. It is part of a larger study which relates personality to school performance, but this paper does not consider school performance. For this reason, the paper is on the borderline for inclusion in our sample of papers reporting research on education. Since in defining the sample we wanted to include basic research which might affect education, as well as practically oriented research, we decided to include this paper. Specifying the "Below average" rating of contribution to educational practice, the judge writes, "Any relation to educational practice is purely inferential, not a direct concern of the study." The author relates his problem to other empirical studies of the relevant variables and seeks to answer unresolved question in the theory of child development. The hypothesis of the study is phrased only on the operational level, however, and not on the theoretical level ("Boys with lower masculinity scores will produce higher self-report anxiety scores").

The last paper which was rated "Above average" with regard to theory and "Below average" in reference to practice is 941. It is one of a series of studies on the non-intellective factors associated with academic achievement among college students. It analyzes conformity to group estimates of physical stimuli as a function of aptitude, performance, and sex under both achievement and affiliation incentive conditions. We agree with the judge's "Above average" rating of contribution to theory because the paper tries to solve some of the unanswered problems arising in earlier work on the relation between social dependence, conformity, and academic achievement. Also, the results are analyzed in terms of the theoretical concepts and relationships specified early in the paper. However, no attempt is made to draw out the implications of this study for educational practice. Therefore, we also agree with the rating of the paper as "Below average" in its contribution to practice.

Characteristics of judges rating these papers

All of the above papers were rated by judges in the same or related specialties. Two of the judges are in schools of education, three are in research institutes. The ones from schools of education rated papers which appeared in education journals or were presented at the 1968 AERA meeting, while the ones in research institutes rated papers which appeared in behavioral science journals. Nothing about the background of the judges suggests why these papers were rated dissonantly with respect to theory and practice.

In short, we have seen that papers rated "Above average" theoretically and "Below average" with regard to practice are generally not directed toward practical purposes. They are not papers which tried to treat a practical problem and failed. Rather, they have a more theoretical or methodological orientation, which they successfully fulfilled. They have been rated "Above average" in contribution to theory, so they are not lacking in substantive worth. In our view, however, even these papers which are rated "Above average" in contribution to theory are not especially theoretical. They do not respecify concepts or create new conceptual frameworks. Also, their hypotheses are usually on the operational level rather than on the theoretical level, as we saw in paper 932. This is very consistent with Bloom's statement that in a given year, only three of the published research studies are crucial and significant for all that follows.[1]

---

[1] Benjamin S. Bloom, "Twenty-five Years of Educational Research," in Educational Research and Contemporary Social Problems, report of the Sixth Annual Convocation on Educational Research, Oct. 18-19, 1965.

Qualitative Analysis of Papers Rated "Below Average" in Terms of Theory and "Above Average" with Respect to Practice

Paper characteristics

Eight papers were rated "Below average" theoretically and "Above average" for practice, and we will examine a random sample of four of them. All of these four examples lack a theoretical source for their hypotheses or problems, and they lack a theoretical framework for the interpretation and explanation of their findings.

The thesis of paper 392 is that a designation such as "culturally deprived" is not only derogatory but inaccurate. Children labeled in this way are in fact diverse. We agree with the judge that this problem is extremely important for educational practice. But the paper makes no effort to account for, or even speculate about, reasons for the diversity in the sample of "culturally deprived" children studied. Is the diversity due to personality variables, family background, or what? The authors make the very important observation that test-related behavior demonstrates even more diversity among children, but they do not mention how test-related behavior might be used to indicate underlying abilities. This study has the potential for questioning theories of testing and measurement, but this issue is never raised. The study calls for refinement of the gross concept "cultural deprivation" which is very much needed in educational practice, but does not specify how the concept can be refined. The judge writes, the paper "adds little to theoretical knowledge but is of practical importance." In our opinion, this is a study that would benefit from a theoretical orientation.

Paper 434 is an analysis of how much biodata adds to the prediction of performance among undergraduates in an education course. Its goal is a practical one, that of discovering student characteristics which lead to success under specified methods of instruction. The judge elaborates the "Below average" rating of theory by checking the statement, "Although the topic represents an important problem, the paper adds little to our understanding of it." We agree with this judgment. The paper says nothing about why certain measurable characteristics might be related to the dependent variable (performance). The paper is concerned with prediction rather than with explanation.

Paper 534 is an examination of the effects of the instructional situation on learning by programmed instruction. Since variations in the instructional situations of programmed instruction are prevalent, we agree with the judge that the study has practical importance. We also agree that the study's contribution to theory is "Below average." There is no theoretical source at all for the hypotheses being tested in the paper. The study shows that students learned more from programmed instruction as an adjunct to other kinds of instruction than groups without programmed instruction did, but it gives no suggestion about why they did. Specifying the "Below average" rating on contribution to theory, the judge notes, "Although the topic represents an important problem, the paper adds little to our understanding of it."

Paper 1203 is perhaps the extreme example of the deviant cases we have been analyzing here. It presents the development of a Career-Orientation Scale for women, which differentiates between career-oriented and non-career-oriented women. The scale consists of a number of interest items. The study provokes many interesting questions, which are never considered. For example, we wonder if certain interests lead to career orientation or if career-orientation fosters certain interests? Also, what are the interests which are related to career orientation? Why does the author think they are related? All of these unanswered questions do not mean the study is without worth. It does make an important practical contribution. The judge writes, "This study is as totally practical and applied as any study could be. It contributes nothing to theoretical knowledge." He did not rate the paper's contribution to theory. We coded his comment as representing a "Below Average" rating, since many other papers with a similar orientation were rated "Below Average" on contribution to theory by other judges.

Characteristics of the judges rating these papers

All of the judges rating these papers were in the same or related specialties as the papers. Two of the judges were in research institutes, and they rated the papers appearing in behavioral science journals, and two were in schools of education and they rated papers given at the AERA meeting. These status similarities give us some reason to believe that characteristics of the papers rather than of the judges are accounting for the differential ratings within papers.

The role of theory in research

The importance of theory in good research emerges more clearly from these qualitative analyses. At best a paper which contributes importantly to theory would provide a new conceptual framework for viewing a problem and possibly a model of the interrelationships of variables. At the least, it would help anticipate relationships between variables and explain results. Without theory, findings may be documented, but they will probably not be understood, and results will only be transferrable when situations are re-created that are identical to the ones in which the original findings were obtained. Theory should increase the generalizability of results.

The Relationship of Theory and Methods Ratings

We will now examine the relationship between theory and research methods ratings.

Quantitative relationship

These two dimensions are related in a way similar to the way theory and practice are related (Table V-14).

TABLE V-14

RELATIONSHIP BETWEEN THEORY AND METHODS RATINGS

Rating of Use of Research Methods	Rating of Contribution to Theory			N
	Above Average	Average	Below Average	
Above average	76	29	9	114
Average	34	48	39	121
Below average	6	28	120	154
	116	105	168	= 389

The ratings of contribution to theory and use of research methods are strongly positively related, although slightly less so than the ratings of theory and practice. (Their Tau Beta is .60 which is significant at the Chi-Square level of .001.)

Qualitative Analysis of Poor Theoretical but Good Methods Papers

Paper characteristics

We will consider a random sample of five of the nine cases in Table V-14 which were rated "Below average" with respect to theory and "Above average" in their use of research methods. We agree with the judges' ratings in all but one case.

Paper 276 studies the effect of reading instruction on the behavior and attitudes toward authority among adolescent delinquent boys. From the population of 2,300 juvenile delinquents handled by a youth agency, 200 boys were found who met certain criteria that were consistent with the objectives of the agency and of the study. Of the 200, volunteers were sought and further criteria were applied, resulting in 45 subjects who were randomly assigned to one of three groups: a reading group, a swimming group, and a control group. The similarity of individuals in each of the groups was measured by analysis of variance on chronological age, reading grade, intelligence quotient, and supervisory period with the youth agency. This analysis showed there were no significant differences between the three groups on those characteristics. People in each of the three groups were given a battery of tests before and after exposure to the programs. We agree with the judge's rating of use of research methods as "Above Average." The judge specifies his "Below Average" rating of theory by checking, "The paper represents an important problem, but adds little to our understanding of it." There is some speculation in the paper about the reasons for certain relationships occurring, but it is based on only a common-sense approach, with no reference at all to theory. For example, the author quotes someone as stating, "Learning difficulties, especially in reading, are among the overwhelming factors contributing to the decompensation of children who have other disturbing problems," but we can't tell if this statement represents an opinion, a theory, or an empirical finding.

Paper 520 is an evaluation of special training in purposive listening. The question raised is: Would already measured differences in purposive listening last over a year? The study uses the experimental design, with one control group given a treatment other than the experimental one, and the third control group given no treatment. The authors present an analysis of variance on the mean scores of the STEP Listening Section to compare the relative position of the three groups. We agree with the judge's evaluation of the paper as "Above Average" in its use of research methods. Specifying the "Below Average" rating of contribution to theory, the judge says, "Extremely limited discussion, no concern for implication." We agree. There is no mention of why these gains might have been made or retained. It is hard to resist saying, "So?" when one finishes reading the paper.

The purpose of paper 586 is to identify instructional program characteristics related to learning effects. The judge writes, "No attempt to provide theoretical framework. May be highly valued by those in limited audience for which it was written. Cryptic style not helpful." We agree. The paper makes no mention of why certain characteristics might be related to learning, or of what the relevant learning theories are. Methodologically the paper is rated as outstanding by the judge, and we agree. The paper reports an exploration of applying multiple regression analysis and factor analysis to the problem being studied. There is a thorough discussion of methodological problems.

Paper 545 was one of the ones rated in common by five judges. Four out of five rated it "Below Average" with respect to theory and three out of five rated it "Above Average" with reference to research methods. The paper describes using reinforcement procedures involving money in conjunction with a token-reinforcement system to maintain appropriate behaviors for learning a complex skill. It does not explain why what was found may have occurred. Some discussion of the relative effects of various rewards on different types of children would have been helpful.

Two judges indicate that the paper does not deal with an important problem, and two think its topic represents an important problem but that the paper adds little to our understanding of it. There is no consistency in the background or situations of the judges which might account for their differences of opinion about why the paper was "Below Average" with respect to theory. If we look at their setting, the ones in schools of education are split on whether or not it represents an important topic. The same holds for their educational background. One trained in the liberal arts for all three degrees thought the paper represented an important problem, and one similarly trained thought it did not deal with an important problem. All but one judge are in specialties related to the topic of the paper (learning, educational psychology) and that one is a specialist in research methods. Perhaps there is something controversial about the topic of the paper which makes it difficult for judges to agree on why they rate it "Below Average" on contribution to theory.

Paper 1267 is an evaluation of a program to improve the attitudes and personalities of high school age mentally retarded students. A random sample of subjects in the target population was drawn. The subjects were classified into a 4 × 2 × 3 factorial design for analysis of variance. The judge indicates that "although the topic represents an important problem, the paper adds little to our understanding of it." The paper does little to explain why subjects further along in the program made better scores on measures of attitudes and personality. It tells us nothing about the nature of the program which may have led to the increases in attitude and personality measures. Perhaps the same factors which were related to higher scores on the dependent variables were related to the student's level of training. This does not seem adequately controlled in the methodology of the study. We would probably rate it "Average" rather than "Above Average" on methods.

In this group of papers with dissonant ratings, we see many of the same explanations for the "Below Average" rating of theory as we saw in pages 129-132. A paper may be well done methodologically, or treat an important practical problem, and yet add little to our understanding of the problem, because it is unrelated to theory.

Characteristics of judges rating these papers

All of the judges rating these papers are specialists in the same sub-fields as the papers. All are located in schools of education. Some of the judges were educated in liberal arts schools, some in schools of education, some in both. All are male and their ages range from 44-69. None of these factors seem to be related to the dissimilar ratings on the two dimensions.

Qualitative Analysis of Good Theoretical but Poor Methods Papers

Paper characteristics

Papers where contribution to theory was rated "Above average" and use of research methods was rated "Below average" are different from those rated in the opposite way. Two out of three (a random sample of the six cases) examined have specific methodological flaws, and the third is on the borderline of our definition of empirical research.

Paper 1159 compares two basic positions in language teaching--the audio-lingual habit theory and the cognitive code-learning theory. We agree with the judge's favorable evaluation of contribution to theory, since the paper tries to compare the relative effectiveness of two theories of language learning. Two experimental groups of three classes each were formed on a random basis. Both received pre-tests and post-tests. Regarding the "Below average" rating of methods, the judge writes, "the study was weakened by making changes in the two treatments that made them non-typical of the methods being studied. Specifically de-emphasis of the language lab in the audio-lingual and extra emphasis of audio work in the cognitive class." We agree with the judge that changes in the two treatments raise serious methodological problems, and warrant the assigned rating.

Paper 484 is an evaluation of the effect of a preschool intervention program. A sample was drawn from available four-year-old disadvantaged children in a school district. Matched experimental and control groups were established on the basis of a number of characteristics the authors considered important. But when we look at their data, the experimental and control groups are not comparable even on the variables on which they were supposed to be matched. Assigning children to control and experimental groups on a random basis would probably have been a better design. The judge comments on the methodology, "Poor instruments [were] used. The methods of data analysis are totally inadequate. I would not let even a Master's student get away with such a shoddy job. It's a real pity since the substantive aspects--both theoretical and applied--are so terribly important." This paper clearly is focused on a crucial question for education. Its problems are ones of execution rather than of topic selection. Consultation with a good research methodologist might have prevented a very important paper from being of little help to those concerned with the same practical problem.

The purpose of paper 386 is to examine (a) some features of the situation of the new university executives (president, vice president, executive deans), (b) their attempts to establish a legitimate basis for authority, and (c) some consequences of these attempts for the conflict that now racks many large campuses. We agree with the judge that the paper is "Above average" in its contribution to theory. The conclusions in the paper are based on (1) a study of the literature on university administration (2) direct observation of several university "administrative councils," and (3) preliminary interviews with a number of large-university administrators. Clearly this is a qualitative study. The judge writes, "This paper was hard to rate on this criterion [methodology]. It is not of the same type as the others in my packet. It is an essay, presenting some very cogent ideas, taking off imaginatively from a review of the literature. I have rated it as 'research,' though it is not. We have no data, but it certainly is not 'incompetent.' Hence my unsure rating." The paper is on the borderline of being empirical research in our definition of the term. If there had not been a systematic effort to collect material for qualitative analysis, we would not have included it in the population.

Characteristics of the judges rating these papers

In an earlier study,[1] we found that people trained in the social sciences and working in a bureau of social research were harsher in their rating of research methods than people in schools of education.

---

[1] [Persell], "Measuring Educational Research Quality."

Two out of three of these papers which were rated "Below average" in their use of research methods were judged so by people located in schools of education, so the rating does not seem to be a function of the judge's location. The other judge is in a regional laboratory. All of the judges are in the same or related specialties as the papers. Two of the judges received all three degrees from liberal arts schools, and the third received a B.A. in the liberal arts and a Ph.D. from a school of education. All are male and their ages range from the late 30's to 69. None of these factors seem related to the ratings.

Relationship of Practice and Methods Ratings

Quantitative Relationship

Like the other dimensions of research quality, ratings of contribution to practice and use of research methods are strongly related (Table V-15), although somewhat less so than ratings on the other dimensions. (Their Tau Beta is .53, which is significant at the Chi-Square level of .10.) Table V-15 also contains the largest number of dissonant ratings.

TABLE V-15

RELATIONSHIP BETWEEN PRACTICE AND METHODS RATING

Rating of Use of Research Methods	Rating of Contribution to Practice			N
	Above Average	Average	Below Average	
Above average	70	30	14	= 114
Average	38	63	20	= 121
Below average	9	40	105	= 154
	117	133	139	= 389

Qualitative Analysis of Good Practical but Poor Methods Papers

Paper characteristics

These papers are typically studying an important problem but have a serious flaw in the research methods employed. Not surprisingly, six out of nine of the papers were rated "Above Average" with respect to theory as well as to practice, so their major difficulty is methodological.

Paper 1191 tried to determine the amount of counseling received from campus sources by subjects in three control groups of a group counseling experiment during a two-term experimental period. We agree with the judge that this is an important practical problem. In the study there were five groups--two experimental and three control. The grade-point-average in the two experimental groups was found to be significantly higher than that in one of the control groups. The judge writes, "Confounding of attrition and change." We agree. Half (93 of 187) of the subjects in all five groups left before the end of the second term. It is not surprising that those remaining had a higher GPA.

Paper 190 has two purposes: to review the performance of foreign students as a check on the standards used for admitting them, and to see what "profile of expectancy" could be established for new foreign students. This study is useful for a specific practical situation, and we agree with the judge who rates it "Above Average" with respect to educational practice. There are virtually no research methods used in the study, however. It is an inventory of the numbers of students from different countries in different categories. The figures are not even presented and compared as percentages. No statistical tests are employed, even for descriptive purposes. The judge checks "poor methods of data collection" and writes, "suspect contaminated or [unequal][1] grading standards."

Papers 386 and 484 have already been described above. We agree that they make an important practical contribution.

Characteristics of the judges rating these papers

All of these judges are in the same or related specialties as the papers. Three out of four of the judges are in schools of education, one is in a liberal arts department. Three received doctoral degrees in the liberal arts, and the fourth received his doctorate in education. All are male and their ages range from the 30's to 69. None of these characteristics seem to be related to the judges' ratings of these papers.

Qualitative Analysis of Good Methods but Poor Practical Papers

Paper characteristics

As we would expect, papers with dissonant ratings in this direction are quite different from the ones we have just been considering. Four out of seven of the papers (a random sample of the fourteen cases) we analyzed seem to have been rated "Below average" with respect to practice because they did not specify or draw out the implications of their findings for educational practice. One study would have contributed to practice if its findings had been positive. Another study seems unrelated to educational practice. It is primarily a methodological study which only incidentally treats an educational topic. Finally, in the case of one paper we do not agree with the "Below average" rating the judge gave its contribution to practice.

---

[1] This word was illegible.

Paper 351 analyzes the effects of learning contexts on students' political behavior. Student subgroups were formed from factor analysis of 999 students' answers to questions about their political behavior. Analysis of variance is used to assess the effect of learning contexts. The effect holds when other characteristics, such as level in school, the amount of time spent in the school situation, the level of father's education, and the individual's own political orientations are held constant. We agree with the judge's evaluation of the methodology as "Above average." This seems to be primarily a methodological study. The author concludes, "This attempt to demonstrate the existence of contextual effects was designed to call attention (once again) to the need for developing empirical techniques to describe the interplay between persons and collectivities." The implications for educational practice are not drawn out, however. The judge says, "Nothing is wrong with the article itself, but the problem is stated and treated so as to have only tangential bearing on 'practice.' The implications for practice, if any, are not drawn out." We think that in view of recent student unrest, the practical significance of this paper may increase.

Paper 98 analyzes teachers' success in managing emotionally disturbed children in regular classrooms. The sampling frame was forty-nine first and second grade classrooms. Each was videotaped for a full day. Each child in the sample was coded for work involvement and deviancy every twelve seconds in a specific academic sub-setting. Also the teachers' behavior and style were coded. Here again the practical implications are not specified. The paper contains a general plea to have the techniques of group management and programming given more emphasis in curriculum for prospective teachers, but nothing in the paper would enable a teacher to alter their behavior on the basis of reading it. The judge specifies the "Below average" rating of contribution to substance by checking, "The topic represents an important problem but it adds little to our understanding of it," and by writing, "faulty interpretation of data."

Papers 520 and 586 have already been described in an earlier section. Paper 520 says nothing about what the special training in purposive listening was, so an educator could not try to apply the techniques. The judge's comment on theory also applies to practice, "extremely limited discussion, no concern for implication."

Paper 586 is similar. It says nothing about the circumstances under which the findings in the paper might make a contribution to educational practice. No reference is made to where or how the learning programs could be used.

Paper 1256 might have made a contribution to educational practice if its results had been positive rather than negative in our view.

Its purpose was to explore the relationship of high need (food, shelter) and low need (love and belonging) contents on arithmetic problem solving. The experimental method was used, with three groups. The findings were analyzed with analysis of variance and factor analysis. The authors found "no statistically significant relationship between need content in arithmetic problem solving and degree of disadvantage." If they had found positive relationships between need contents and achievement for children of different backgrounds, this might have made an important practical contribution to education. Since the results were negative, the practical contribution is negligible.[1] The judge says, "[the paper] deals with an important problem but the investigator studied it unimaginatively."

Paper 585, as we have already noted, is primarily a methodological study.

Paper 276 has already been described. It treats the effect of reading instruction on juvenile delinquents. The judge rated it lower on contribution to practice than we would have.

Characteristics of judges rating these papers

One judge rated two of these papers, so there are six different judges. Two of them are in research institutes, four are in schools of education. All are in the same or related substantive specialties as the papers. Four had degrees in both liberal arts and education and two received all three degrees in the liberal arts. All are male, with their ages ranging from 34 to 69. There seems to be no apparent relationship between these characteristics of the judges and their dissonant ratings.

---

[1] Negative findings are frequently underutilized in social research. They might be used more than they are by practitioners to avoid employing an approach that had been tried and found wanting.

General Characteristics of the Papers with Dissimilar Ratings

Fourteen out of thirty-four of the papers with dissimilar ratings are AERA papers, although AERA papers represent only one quarter of all papers rated. However, of the fifteen papers rated "Below average" on research methods and "Above average" on theory or practice, only three were AERA papers. So, although AERA papers are disproportionately represented in this group of disparate ratings as a whole, they are less likely to be deficient in their use of research methods. They are more likely to be rated "Below average" with respect to theory or practice.

General Characteristics of Judges Giving Dissonant Ratings

We have not seen any apparent relationships between characteristics of judges (e.g., location, field of degrees, specialty, age) and particular types of dissonant ratings. However, certain judges seem to be generally more likely to give a disparate rating, although they may not have a tendency to rate one dimension higher or lower than another. Twenty-one of forty judges gave one or more dissimilar ratings, and of these, thirteen gave two or more (eight gave two, two gave three, and three gave four). Of these thirteen, all but four received all their degrees in the behavioral sciences, and those four received one or two of their three degrees in the behavioral sciences. All but two of the thirteen are now in a school of education or an educational research institute. Thus, most were trained in the behavioral sciences, but work in schools of education. It may be that a certain creative tension exists between the educational background of these judges and their current setting which increases the chance that they will rate each dimension of a research paper independently of the other dimensions. This does not mean that the variation in the three aspects of quality does not exist in the papers themselves (of the ones we analyzed, we almost always agreed with the judges' ratings). Rather it may mean that these people are particularly sensitive to variations in quality on different dimensions. They may be less likely to rate all three dimensions in a similar fashion.

The Relationship of All Three Dimensions of the Rating Form

Having seen the relationship between each of the dimensions of the Rating Form, and analyzed some of the deviant cases in these relationships, we will now look at the relationship of all three dimensions simultaneously. There is a strong positive relationship among all three dimensions of research quality. (Their average Tau Beta is .59.) (Table V-16) Clearly the research papers in this population were generally rated as being similar on all three dimensions of quality. The papers are not particularly good or particularly poor on one dimension compared to the others. Yet there is enough variation in the

TABLE V-16

RELATIONSHIP OF ALL THREE DIMENSIONS
OF THE RATING FORM
(N=389)

Rating of Contribution to Practice	Rating of Contribution to Theory								
	Above Average			Average			Below Average		
	Rating of Use of Research Methods								
	Above Average	Average	Below Average	Above Average	Average	Below Average	Above Average	Average	Below Average
Above average	57	11	2	22	12	4	4	3	2
Average	14	13	3	12	31	20	2	17	21
Below average	5	5	4	0	5	15	0	8	97

evaluations of the different dimensions so it will be interesting to keep the three dimensions separate for most of our analysis.

Nearly twice as many papers are rated "Below average" on all three dimensions as are rated "Above average," 97 papers compared to 57 papers. If we look at only those papers rated "Best" (1) and "Incompetent" (5) on all three dimensions, we find that this proportion is even greater. Only four papers were rated "Best" on all three dimensions, compared to 20 papers rated "Incompetent" on all dimensions. These figures suggest that general estimates of the low quality of educational research are justified. The relatively small amount of outstanding work does not counterbalance the large amount of poor work. The large proportion of poor educational research emphasizes the importance of discovering what affects variation in research quality. This is the problem we address in Part II of this volume.

Summary

The distribution of research quality in this population reveals that a considerable amount, 43 per cent, of published research on education in 1967-68 was rated "Below average" or "Incompetent" with respect to one or more of the three dimensions of quality. This substantiates what many observers of educational research have sensed for some time. It also means that research on education faces some grave problems. If it is to add to knowledge and ultimately contribute to improving education, it needs to be better.

In this chapter we have seen that the three dimensions of research quality are highly interrelated. Qualitative analysis of papers which were rated high on one dimension and low on another illustrated the importance of theory in explaining findings. These analyses also revealed that papers rated "Below average" on the use of research methods and high on other aspects generally contained specific flaws in their sample, design, or instruments. Finally, papers which were rated "Below average" with respect to practice seem to be those which failed to draw out the implications of their findings for educational practice.

We found no apparent relationship between the background and location of judges and the direction of their dissonant ratings. There does seem to be some relationship between the judges' educational background and current location and their tendency to give a dissonant rating. Specifically, those trained in the behavioral sciences but working in schools of education seem most likely to give discordant ratings to the different dimensions.

PART TWO

EXPLAINING RESEARCH QUALITY

CHAPTER VI

ROLE SOCIALIZATION AND QUALITY

Numerous studies have investigated the way training affects research outcomes, but they differ in several important respects from the one reported here. First, they generally measure research outcomes in terms of productivity rather than quality. Investigators of productivity in this and related fields include Ben-David, Buswell, Lingwood, and Worthen.[1] Another difference between this study and the ones mentioned above is that they are generally not conceptualized in theoretical terms. Except for Ben-David's work, which identifies the general conditions which precede or accompany increases in research productivity, the other studies report empirical findings, but offer little explanation of why they occur. Carefully conducted empirical studies such as these are important for adding to our knowledge of the phenomenon, however, and they have been helpful to us.

The study which comes closest, in our opinion, to relating training and career experiences to career outcomes in a theoretical framework is Wilder's study of reading experts.[2] The career outcome which Wilder studied was research involvement and orientation. Unlike productivity, research experience is strongly related to quality (Table VI-13). Wilder's study also offers a theoretical explanation of why reading experts are not more research oriented, namely because the applied science of reading has not been institutionalized. But Wilder studied only reading experts, so we do not know if hie findings apply to all types of educational research. Also, Wilder did not measure quality. One of the few studies which examines quality is Cole's investigation of physics,[1] but he does not consider how professional socialization affects quality.

---

[1] Joseph Ben-David, "Scientific Productivity and Academic Organization in 19th Century Medicine," in Norman Kaplan, Science and Society, pp. 39-61; Buswell, et al., Training for Educational Research; Lingwood, "Interpersonal Communication"; Worthen, "Impact of Research Assistantship Experience."

[2] Wilder, "Reading Experts."

The study reported in this volume differs from previous work in two other respects as well. We consider training within the theoretical framework of role socialization.[2] Also, we see socialization as one of several major classes of variables that affect quality, and we consider the authors' background, their organizational settings, and the larger context in which they work as well.

In this chapter we will focus on the following five elements of role socialization:

(1) type of doctorate

(2) practice-oriented socialization

(3) graduate research experience

---

[1] Jonathan R. Cole, "The Social Structure of Science: A Study of the Reward and Communications Systems of Modern Physics" (unpublished Ph.D. dissertation, Columbia University, 1969).

[2] Suggested by Basil J. Sherlock and Richard T. Morris, "The Evolution of the Professional: A Paradigm," Sociological Inquiry, XXXVII (Winter, 1967), 27-46.

(4) career research experience

(5) research orientation

We consider these five because our role socialization perspective and previous empirical findings suggest they will be key concepts for explaining quality. We will consider these five elements in turn, discussing how they were measured and how they are related to quality. Then, in the second section of this chapter, we develop a model of hypothesized interrelationships between these elements and quality.

Elements of Role Socialization

Type of Doctorate

We expect that the type of doctorate[1] an author received affects the nature of his training, career experience, orientation, and the quality of his work. We measured type of doctorate in two ways:

(1) We asked respondents to indicate the major field (see Question 3.1 in Questionnaire, Appendix II-5) of each of their degrees. These doctoral majors were coded into five broad categories: education, education fields (e.g., science education), psychology, social science, and other, using Bargar's revised list of subfields.[2]

(2) We also asked authors about each of their degrees, "Was degree in education?" (see Question 3.1 in Questionnaire). By this is meant was the degree obtained in a school or department of education.

---

[1] All but 11 per cent of this sample have doctorates. Where our analysis includes type of doctorate as a variable, we consider only those with doctorates, but in other analyses we consider those without doctorates as well.

[2] Bargar, et al., National Register of Educational Researchers, pp. 12-13.

As we would expect, there is a very strong relationship between the two measures (Table VI-1). Because there is so much congruence between major field and place of doctorate, we could use either one to measure type of degree. We decided to use whether their degree was from a school of education or not, because of the important contextual properties of schools of education.[1] But it should be remembered throughout our discussion that type of doctorate really refers to both the context and the field of an author's doctoral training. When we compare authors with regard to their type of doctorate, we speak of education doctorates (those in column (1) of Table VI-1) and behavioral scientists (those in column (2) of Table VI-1).[2]

Some people would argue that type of doctorate should not affect research quality, since education is a discipline just as the behavioral sciences are. But there are substantive and career reasons for expecting behavioral scientists to do better research than education doctorates. We believe the level of theoretical development varies by field. Theoretical development can be measured on at least two dimensions: consensus and "hardness." As Smelser notes, a discipline achieving scientific maturity

> . . . more nearly attains consensus on the scientific problems to be posed, the relevant independent variables, a theoretical and philosophical perspective, and appropriate research methods. Simultaneously, it witnesses a decline of distinctive schools;

---

[1] For example, Hagstrom, "Educational Researchers"; Sieber and Lazarsfeld, The Organization of Educational Research.

[2] Using the term type of doctorate should not be confused with the distinction between Ph.D. and Ed.D. degree. While all behavioral science doctorates have Ph.D.'s, 62 per cent of education doctorates also have Ph.D.'s and 38 per cent of educational doctorates have Ed.D.'s.

TABLE VI-1

DOCTORAL MAJOR BY WHETHER DOCTORATE WAS IN EDUCATION OR NOT

Doctoral Major	Was Doctorate in Education?	
	Yes (1)	No (2)
Education	423	13
Education Fields	58	5
Psychology	6	152
Social Science	2	56
Other	3	24
No Doctorate	12	7
	504	257

> a decline in the quantity of polemic about the "nature" of the field and the value of different "approaches" to the field, a decline in propaganda, proselytization, and defensiveness; and an increase in discussion of findings in relation to accepted criteria of validation.[1]

Thus, the more developed the field, the more agreement on how problems should be approached and how results should be evaluated. The relative "hardness" of a science is related to the "tightness of integration of their various bodies of knowledge," as Storer suggests.[2] To measure the tightness of integration, he suggests the frequency with which mathematics is used in the different sciences. Before mathematics can be used in a field, of course, the field needs to produce general statements which can be phrased in propositional form. On both of these measures we think the behavioral sciences reveal greater theoretical development than education. This is a substantive reason for expecting behavioral scientists to do better research than education doctorates.

In addition, behavioral science doctorates are less likely to have had practice-oriented socialization, they are more likely to have research experience during graduate school and more likely to have research experience during their careers (see Tables 1 through 3 in Appendix VI-1). This developmental sequence associated with type of degree is a career-related reason for expecting behavioral scientists to do better research than education doctorates.

---

[1] Neil J. Smelser, "Sociology and the Other Social Sciences," in The Uses of Sociology, ed. by Paul F. Lazarsfeld, William H. Sewell, and Harold L. Wilensky (New York: Basic Books, Inc., 1967), pp. 7-8.

[2] Norman W. Storer, "The Hard Sciences and the Soft: Some Sociological Observations," Bulletin of the Medical Library Association, LV (January, 1967), 75-84.

degree In line with our expectations, behavioral scientists do better research than education doctorates on all three dimensions of quality--theory, practice, and methods (Table VI-2). Overall, 41 per cent of behavioral scientists are doing good research compared to 27 per cent of education doctorates. In the next section we will try to understand which of the above characteristics of type of doctorate--differences in theoretical development or differences in socialization--seem to explain why behavioral scientists do better research.

Practice-oriented Socialization

Previous empirical studies have found several types of experience that are negatively related to research outcomes, but they do not really explain why these experiences affect research adversely. Let us consider three sets of data. Buswell found that elementary or secondary school teaching experience is negatively related to later research productivity.[1] Wilder found it is negatively related to research orientation and behavior,[2] and we find it is negatively related to research quality (Table VI-3). Of those with no teaching experience, 39 per cent do good research compared to 19 per cent of those with five or more years of teaching experience.[3]

---

[1] Buswell, et al., Training for Educational Research, p. 16.

[2] Wilder, "Reading Experts," p. 116.

[3] Nor are these differences in quality explained by different levels of undergraduate performance among those with more or less teaching experience. High performers are just as likely to have raight as not.

TABLE VI-2

RESEARCH QUALITY BY TYPE OF DOCTORATE

Research Quality		Type of Doctorate			
		Behavioral Sciences		Education	
Contribution to theory:					
Best[a]	good	7%	41%	3	27%
Better than average		34		24	
Average		26		29	
Below average		24		26	
Incompetent		10		18	
		101%		100%	
		(98)		(201)	
Contribution to practice:					
Best	good	10	39%	5	27%
Better than average		29		22	
Average		42		37	
Below average		14		25	
Incompetant		5		12	
		100%		101%	
		(98)		(201)	
Use of research methods:					
Best	good	8	40%	5	29%
Better than average		32		24	
Average		29		35	
Below average		22		23	
Incompetant		9		14	
		100%		101%	
		(98)		(200)	

[a] In this case only we are retaining the five-point rating scale, since it shows the differences between the fields more clearly.

TABLE VI-3

RESEARCH QUALITY BY SCHOOL TEACHING EXPERIENCE

Research Quality	Number of Years Spent Teaching Elementary or Secondary School		
	None	One-Four Years	Five or More Years
Contribution to theory:			
Above average	41%	35	14
Average	30	21	29
Below average	29	45	57
	100%	101%	100%
	(153)	(87)	(112)
Contribution to practice:			
Above average	37	31	21
Average	37	31	36
Below average	26	38	43
	100%	100%	100%
	(153)	(87)	(112)
Use of research methods:			
Above average	40	26	22
Average	33	36	30
Below average	28	38	48
	100%	100%	100%
	(152)	(87)	(112)

In addition, in separate studies Brown, Buswell, and Millikan found that Ph.D. recipients are more productive researchers than Ed.D.'s.[1] Wilder found that reading experts with Ph.D.'s differ markedly from Ed.D.'s in terms of their research activity, career patterns, values, type and place of work, even when age, sex, community of origin, religion, and school attended are held constant.[2] We find that Ph.D.'s do better research than E.D.'s (Table VI-4). Among Ph.D. recipients, 32 per cent do good research compared to 25 per cent of Ed.D. recipients.[3]

Finally, Wilder found that professional education courses decrease research orientation, and we find that they have a negative linear relationship to better research (Table VI-5). Among those who never took a professional education course, 43 per cent are doing good research, compared to only 25 per cent of those who took more than three courses.

---

[1] L. D. Brown, Doctoral Graduates in Education: An Inquiry into Their Motives, Aspirations, and Perceptions of the Program, Cooperative Research Project No. S-240, Office of Education, U.S. Dept. of Health, Education, and Welfare (Washington, D.C.: Government Printing Office, 1966); Buswell, et al., Training for Educational Research; Nancy M. Millikan, The Development of Professional Personnel in Educational Research, Cooperative Research Project No. S-487-64 (New York: Teachers College, Columbia University, 1967).

[2] Wilder, "Reading Experts," Ch. VI.

[3] These differences are enhanced when we control for type of degree. Ph.D.'s in the behavioral sciences do even better research than Ph.D.'s in education, with Ed.D.'s least likely to do good research. The percentages doing good research are, respectively, 38 per cent, 29 per cent, and 25 per cent.

Because of the wide variation in the curriculum of Ed.D. programs, it is possible that a particular Ed.D. program provided more research training than a particular Ph.D. program. The results reported here should be viewed with this possible qualification in mind.

TABLE VI-4

RESEARCH QUALITY BY NAME OF DEGREE

Research Quality	Name of Degree	
	Ph.D.	Ed.D.
Contribution to theory:		
Above average	34	21
Average	26	30
Below average	39	49
	99%	100%
	(224)	(86)
Contribution to practice:		
Above average	29	30
Average	41	30
Below average	30	40
	100%	100%
	(224)	(86)
Use of research methods:		
Above average	34	24
Average	31	34
Below average	35	42
	100%	100%
	(223)	(86)

TABLE VI-5

RESEARCH QUALITY BY NUMBER OF PROFESSIONAL EDUCATION COURSES

Research Quality	Number of Professional Education Courses Taken in Graduate School			
	None	One-Two	Three	More Than Three
Contribution to theory:				
Above average	45%	42	35	24
Average	33	23	26	26
Below average	22	36	39	51
	100%	101%	100%	101%
	(78)	(31)	(23)	(212)
Contribution to practice:				
Above average	40	39	35	26
Average	42	32	35	33
Below average	18	29	30	41
	100%	100%	100%	100%
	(78)	(31)	(23)	(212)
Use of research methods:				
Above average	44	45	30	24
Average	31	32	39	33
Below average	25	23	30	43
	100%	100%	99%	100%
	(77)	(31)	(23)	(212)

Why do these experiences--school teaching, professional education courses, and the Ed.D. degree--have such an adverse effect on research? Let us examine one of them in more detail. If we compare the role of school teacher and that of researcher we find that the norms and expectations attendant upon the two roles are quite different. Typically, the school teacher works in a hierarchical system of authority. Rewards generally go to those who contribute to the smooth functioning of the system, and school teachers usually find friends and obtain social support from other teachers. Furthermore, in the classroom, teachers are exposed to the immediate needs of individual children. At the same time, an elementary or secondary school teacher is more likely to be isolated from the theoretical and methodological developments of a substantive field, as well as from the social and normative supports of others in a substantive field.

In contrast, the researcher more often works in a less hierarchical authority system. He generally seeks recognition and social support from other researchers in his substantive field, and rewards are assumed to go to work which contributes to knowledge and understanding.[1] Finally, the researcher is not faced with the daily needs and problems of individual children. As a result, a researcher has different orientations, different friends, and he is rewarded for different things than is the school teacher.

In view of these differences we would expect authors who had taught elementary or secondary school to be less research oriented than those who had not taight. Like school teaching, professional education courses and the Ed.D. degree may be seen as diverting authors away from research and directing them toward the practice of education. In brief, these three experiences may affect research adversely because of their effect on the authors' orientations. We will consider this interpretation further in the next section.

---

[1] This assumption about the operation of the reward system in educational research will be analyzed in Chapter IX.

Since these three experiences are related empirically to each other, we combined them into an index of practice-oriented socialization.[1] (See Appendix VI-2 for the matrix of correlations among the items and for a description of how the index was formed.) Not surprisingly, practice-oriented socialization has a strong negative linear relationship to quality (Table VI-6). Among those with no practice-oriented socialization, 42 per cent are doing good research compared to 20 per cent of those with more such experience. Contributions to theory benefits most from the lack of practice-oriented socialization; 46 per cent of those without it are contributing to theory compared to 14 per cent of those with such socialization.

Graduate Research Experience

Buswell and Wilder found that research experience during graduate school is importantly related to later research productivity or research orientation.[2] We measured two types of graduate research experience.

---

[1] We use the term socialization here because the index includes both training and experience.

[2] Buswell, et al., Training for Educational Research; Wilder, "Reading Experts."

TABLE VI-6

RESEARCH QUALITY BY PRACTICE-ORIENTED SOCIALIZATION

Research Quality	Amount of Practice-Oriented Socialization			
	None	Less	Some	More
Contribution to theory:				
Above average	46%	33	34	14
Average	30	33	22	28
Below average	25	33	44	58
	101%	99%	100%	100%
	(61)	(66)	(132)	(85)
Contribution to practice:				
Above average	36	35	30	26
Average	46	35	36	26
Below average	18	30	34	48
	100%	100%	100%	100%
	(61)	(66)	(132)	(85)
Use of research methods:				
Above average	45	36	29	20
Average	32	32	33	33
Below average	23	32	38	47
	100%	100%	100%	100%
	(60)	(66)	(132)	(85)

First we asked authors: "While in graduate school, did you ever work as a research assistant?" Work as a research assistant is assumed to mean either that an author helped a university professor with his research or that he assisted on a research project which may or may not have been directed by a faculty member. Separately we asked, "As a graduate student did you ever work in a research center or bureau within the university?" When we cross-tabulated the responses to these two questions, we have the following combinations[1] of graduate research experience:

(1) Those who worked both as research assistance and in research centers (198 respondents).

(2) Those who worked as research assistants but did not work in research centers (295 respondents).

(3) Those who had neither form of graduate research experience (200 respondents).

These three combinations may be considered an index of graduate research experience. Those in the first group were exposed to research both through the nature of their work and through the context in which they worked, whereas those in the second group did not work in a context devoted to research and may have less research experience than those in the first group. Those who did neither have less experience still, and may have done no research other than what was required for their courses. courses.We expect people with both types of graduate research experience to do better research than those with one or none and they do (Table VI-7). Of those with both experiences 38 per cent do better research, compared to 25 per cent of those with neither type of experience. We expect people with both experiences to do less poor research[1] as well, but this occurs only with respect to research methods. People who worked as research assistants but *not* in research centers are least likely to do poor research with respect to theory and practice. This slightly anomalous finding may be due to the varied quality of research centers.

---

[1] There were also twenty-four authors among our 901 respondents who worked in a research center while a graduate student, but not as a research assistant. They may have been graduate student research project directors. There are too few to consider in our analysis above, but it is worth noting that of the eleven whose papers were rated, 64 per cent made an above average contribution to theory, 55 per cent made an above average contribution to practice, and 55 per cent were above average with respect to their use of research methods.

An indirect way to measure the quality of research centers is by measuring the prestige of the department in which they are located. Using the reputational ratings of graduate departments reported by Sieber[2] and Cartter,[3] we constructed a rough index of departmental prestige. (See Appendix VI-3 for a description of how this index was formed.) Authors who had both types of research experience in prestigious departments do better research on all three dimensions than people with neither experience; 41 per cent compared to 23 per cent are doing

---

[1] Usually we present only the upper third of the quality ratings in three variable tables because the lower third is consistent with it, i.e., those doing the best research are least likely to do poor research. In this case we present the lower third as well, because it is not consistent with the upper part.

[2] Sieber, *Organization of Educational Research*, pp. 273-80.

[3] Cartter, *An Assessment of Quality in Graduate Education*.

TABLE VI-7

RESEARCH QUALITY BY GRADUATE RESEARCH EXPERIENCE

Research Quality	Graduate Research Experience		
	Both*	Research Assistant Only	Neither
Contribution to theory:			
Above average	37%	33%	27%
Average	25	38	22
Below average	39	29	52
	101%	100%	101%
	(85)	(96)	(83)
Contribution to practice:			
Above average	33%	35%	24%
Average	29	39	40
Below average	38	26	36
	100%	100%	100%
	(85)	(96)	(83)
Use of research methods:			
Above average	43%	31%	24%
Average	31	37	36
Below average	26	32	40
	100%	100%	100%
	(84)	(96)	(83)

*Both assistant and in research center.

good research (Table VI-8). The relationship is contingent, however, upon high departmental prestige. That is, there is no relationship between graduate research experience and better research when the experience was obtained in less prestigious departments. In short, graduate research experience is related to better research only when it occurs in prestigious departments.

When we look at poor research, however, the findings are not quite parallel to those concerning better research. Authors who worked as research assistants but not in research centers in prestigious departments are least likely to do poor research, when compared to those with both experiences and those with neither. This finding suggests that working in a research center even in a prestigious department does not prevent poor research. It may be that quality control is transmitted more effectively under personalized conditions. Perhaps work in a research center is more likely to occur in large, impersonal universities, while working as a research assistant outside of a bureau occurs under more personalized circumstances. This is a problem that warrants further investigation.

Career Research Experience

If we view socialization for the research role as a process that does not end with graduate training but continues throughout one's career, then we would expect career research experience to affect quality. To illustrate, assuming two authors were equally trained we would expect the one with more career research experience to do better research than the one with little research experience after graduate

TABLE VI-8

RESEARCH QUALITY BY GRADUATE RESEARCH EXPERIENCE BY DOCTORAL DEPARTMENT PRESTIGE

Research Quality	Doctoral Department Prestige					
	High			Other		
	Type of Graduate Research Experience					
	Both*	Ass't but not in Bureau	Neither	Both*	Ass't but not in Bureau	Neither
Contribution to theory:						
Above average	39	38	24	29	31	32
Average	27	33	28	26	39	17
Below average	34	29	48	46	30	51
	100%	100%	100%	100%	100%	100%
	(44)	(24)	(29)	(35)	(64)	(41)
Contribution to practice:						
Above average	39	33	21	20	36	29
Average	27	38	52	37	39	42
Below average	34	29	28	43	25	29
	100%	100%	101%	100%	100%	100%
	(44)	(24)	(29)	(35)	(64)	(41)
Use of research methods:						
Above average	46	38	24	35	28	24
Average	36	42	45	29	33	32
Below average	18	21	31	35	39	43
	100%	101%	100%	99%	100%	99%
	(44)	(24)	(29)	(34)	(64)	(37)

*Both assistant and in research center.

school. We measured career research experience with four indicators: time division, six months or more doing research, average rank of research in one's career and research center staff membership.

Time division

We asked respondents, "At the time of this research how did you divide your professional time?" The way people spend their professional time is one indicator of their career experience. Wilder found that only 15 per cent of reading experts spent more than 111per cent of their time on research,[1] and Bargar found that only 34 per cent of educational "researchers" spent more than 30 per cent of their time doing research.[2] While the authors we are studying generally spend more time on research than did those in Bargar's or Wilder's study (72 per cent in this sample spent more than 20 per cent of their time on research at the time they wrote their paper), there are still very few full time researchers. Only 9 per cent of authors in this sample were full time researchers when they wrote their papers. We expect that the more time spent doing research, the better the quality. As expected, the percentage of time an author spent on research when he wrote the paper has a positive linear relationship to better research (Table VI-9). Of those spending one-fifth time or less on research, 25 per cent are doing good work, compared to 39 per cent of those spending more than four-fifths time doing research.

---

[1] Wilder, "The Reading Experts," p. 84.

[2] Bargar, et al., National Register of Educational Researchers, p. 84.

TABLE VI-9

RESEARCH QUALITY BY PERCENTAGE OF TIME SPENT ON RESEARCH AT TIME PAPER WAS WRITTEN

Research Quality	Percentage of Time Spent on Research at Time Paper was Written			
	0-20%	21-50%	51-80%	81-100%
Contribution to theory:				
Above average	22%	32	47	32
Average	33	26	21	39
Below average	45	43	33	29
	100%	101%	101%	100%
	(95)	(143)	(43)	(28)
Contribution to practice:				
Above average	27	31	37	43
Average	39	34	35	32
Below average	34	35	28	25
	100%	100%	100%	100%
	(95)	(143)	(43)	(28)
Use of research methods:				
Above average	26	33	31	43
Average	34	32	33	39
Below average	40	35	36	18
	100%	100%	100%	100%
	(95)	(143)	(42)	(28)

Six Months or More Doing Research

General research experience throughout one's career was measured by asking, "Other than while you worked on your master's or doctoral thesis, has there ever been a period of six months or more when research was your primary activity?" Forty-eight per cent of respondents indicate that there was. This measure indicates a period of intense research experience, even if for a relatively short period of time, and therefore we would expect it to be related to quality. Thirty-nine per cent of people who spent six months or more doing research do better work, compared to 24 per cent of those who have not spent six months on research (Table VI-10).

Average rank of research in one's career

The importance of research compared to other career activities throughout an author's career was measured by computing the average rank of research as an activity throughout his career.[1] Many more authors rank teaching as relatively high in their careers than rank research high (high being an average rank of 2.00 to 3.49 throughout their careers). For 43 per cent, teaching is ranked high in their careers compared to only 18 per cent for whom research is ranked high.

The average rank of research in one's career is positively related to better research (Table VI-11). People for whom research has

---

[1] This average rank of research is only an approximate measure since we did not weight these ranks by the number of years in each position that a given activity had a particular rank. This measure was obtained from data in Question 2.12 in the Questionnaire (see Appendix II-5).

TABLE VI-10

RESEARCH QUALITY BY SPENDING SIX MONTHS OR MORE WHEN RESEARCH WAS THE PRIMARY ACTIVITY

Research Quality	Spent Six Months or More with Research as Primary Activity	
	Yes	No
Contribution to theory:		
Above average	39%	24
Average	26	29
Below average	35	48
	100%	101%
	(167)	(182)
Contribution to practice:		
Above average	37	25
Average	38	33
Below average	25	42
	100%	100%
	(167)	(182)
Use of research methods:		
Above average	40	23
Average	34	31
Below average	27	46
	101%	100%
	(166)	(182)

TABLE VI-11

RESEARCH QUALITY BY AVERAGE RANK OF
RESEARCH IN AUTHOR'S CAREER

Research Quality	Average Rank of Research in Author's Career						
	Always 1st to 1.99	2.00-2.49	2.50-2.99	3.00-3.49	3.50-3.00	Always Fourth	No Research in Some Jobs
Contribution to theory:							
Above average	0	47%	38	27	19	29	21
Average	0	19	33	28	25	35	23
Below average	0	33	30	45	56	35	57
		99%	101%	100%	100%	99%	101%
	0	(57)	(40)	(96)	(32)	(51)	(44)
Contribution to practice:							
Above average	0	40	28	31	16	28	30
Average	0	28	50	32	41	35	36
Below average	0	32	23	37	44	37	34
		100%	101%	100%	101%	100%	100%
	0	(57)	(40)	(96)	(32)	(51)	(44)
Use of research methods:							
Above average	0	41	38	30	19	26	23
Average	0	29	38	37	31	33	27
Below average	0	30	25	33	50	41	50
		100%	101%	100%	100%	100%	100%
	0	(56)	(40)	(96)	(32)	(51)	(44)

a higher rank in their careers are doing better research than people for whom research has a lower rank (43 per cent compared to 25 per cent are doing good research).

Research center staff membership

Besides time expenditures, research experience can be measured by staff membership in a research center at the time the paper was written (see Question 1.11 in Questionnaire Appendix II-5). Research experience in a research center was helpful for graduate students, so we expect it to be helpful in the course of one's career, for many of the same reasons. Being a staff member of a research center increases the chance of regular intellectual interaction and the chance of finding colleagues with complementary talents and skills, as well as providing physical facilities for research. Therefore we expect that being a staff member of a research center will be related to doing better research. As expected, research staff members are doing better research (Table VI-12); 40 per cent of them do good research compared to 28 per cent of non-staff.

These four indicators of career research experience--percentage of time spent on research, six months or more spent on research during one's career, average rank of research in one's career, and research center staff membership--are all positively related to each other. Therefore we combined all but the average rank of research in the author's career into an index of career research experience. The fact that the average rank of research is highly related to the index provides an independent validation of the index. (See Appendix VI-4 for

TABLE VI-12

RESEARCH QUALITY BY STAFF MEMBERSHIP IN RESEARCH CENTER

Research Quality	Was Author a Staff Member of a Research Center When He Wrote Paper?	
	Yes	No
Contribution to theory:		
Above average	37	29
Average	26	28
Below average	37	43
	99%	100%
	(86)	(255)
Contribution to practice:		
Above average	44	26
Average	29	38
Below average	27	36
	100%	100%
	(86)	(255)
Use of research methods:		
Above average	39	28
Average	33	33
Below average	28	39
	100%	100%
	(85)	(255)

1810

the intercorrelation matrix of these items and for a description of how the index was constructed.)

We would be very surprised if people with more research experience were not doing better work. As we anticipated, people with more research experience in their careers are doing considerably better research than those with less experience (Table VI-13). Among the experienced, 43 per cent are doing good research compared to 24 per cent of the less experienced.

Research Orientation

We measured the authors' orientation toward research at the time he completed the questionnaire, by asking respondents to rank professional activities in terms of their personal preference for them (see Question 2.3, Appendix II-5). Authors who prefer research first as a professional activity were considered to be research-oriented. Twenty-two per cent of the authors prefer research, compared to 34 per cent who prefer teaching, 42 per cent who prefer service, and 3 per cent who prefer some other activity.[1] We hypothesize that research orientation is both an outcome of prior experiences that affect research favorably, and an indicator of whether an author is currently attracted to research. (These hypothesized relationships will be explored in the next section.) Therefore we expect that research

---

[1] We defined service as "developing or implementing new curriculum or action programs, school surveys, test administration or scoring, workshops." Also included in this service category are the responses "in-service training of teachers" and "translating research into educational practice."

TABLE VI-13

RESEARCH QUALITY BY CAREER RESEARCH EXPERIENCE

Research Quality	Career Research Experience		
	High	Medium	Low
Contribution to theory:			
Above average	45%	28	23
Average	24	29	31
Below average	31	44	46
	100%	101%	100%
	(91)	(87)	(124)
Contribution to practice:			
Above average	44	30	24
Average	31	43	35
Below average	25	28	41
	100%	101%	100%
	(91)	(87)	(124)
Use of research methods:			
Above average	40	33	24
Average	32	39	30
Below average	28	28	46
	100%	100%	100%
	(90)	(87)	(124)

oriented authors will do better research than authors with other orientations. In line with our expectations, research oriented authors are doing better research than either teaching oriented or service oriented ones (Table VI-14). Of the research oriented, 44 per cent are doing good research, compared to 26 per cent of the teaching oriented and 17 per cent of the service oriented.

The Interrelation of the Elements

In the preceding discussion we have provided some preliminary indications of how the elements of role socialization are related to each other and to quality. In Figure VI-1 we present these relationships in a model of hypothesized associations between socialization and quality.

We just saw that research orientation is strongly related to better research. Is it also true, as we suggested above, that research orientation is affected by type of doctorate, practice-oriented socialization, and career research experience? As suggested, each of these variables is related to research orientation (Table VI-15). Behavioral science doctorates, authors with less practice-oriented socialization, and those with more career research experience are more research oriented than their counterparts. Rather than examining each of these two-way associations individually, however, we consider the joint association of all three in relation to research orientation, in order to assess their relative importance. By doing this we can see if each of these experiences contributes independently to an author's research orientation.

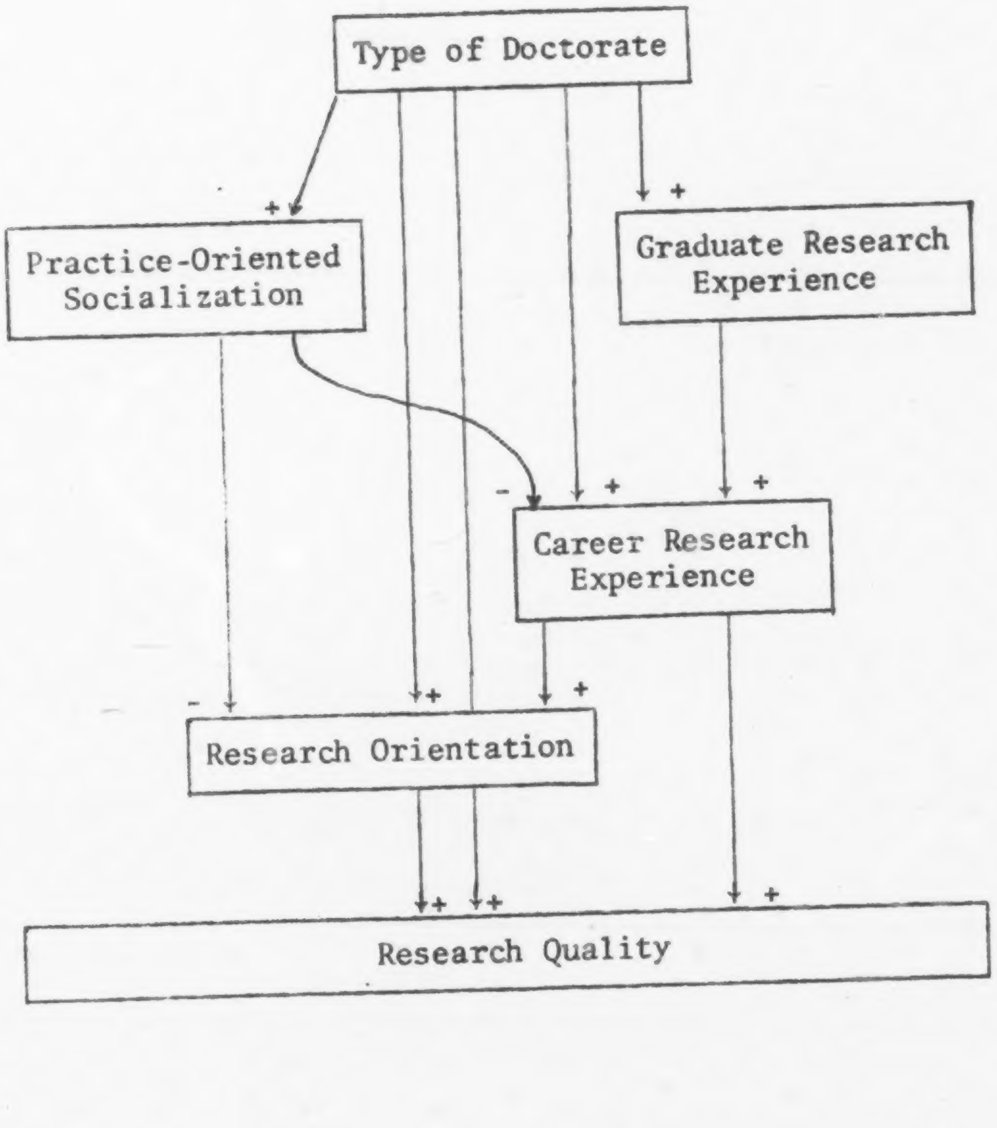
TABLE VI-14

RESEARCH QUALITY BY ORIENTATION

Research Quality	Orientation			
	Research	Teaching or Administration	Service	Other
Contribution to theory:				
Above average	45%	25	12	21
Average	26	28	29	33
Below average	29	47	59	46
	100%	100%	100%	100%
	(124)	(156)	(41)	(24)
Contribution to practice:				
Above average	40	26	17	33
Average	36	36	37	33
Below average	23	39	46	33
	99%	101%	100%	99%
	(124)	(156)	(41)	(24)
Use of research methods:				
Above average	47	22	22	21
Average	31	34	32	38
Below average	22	44	46	42
	100%	100%	100%	101%
	(123)	(156)	(41)	(24)

FIGURE VI-1

HYPOTHESIZED MODEL OF HOW THE ELEMENTS OF ROLE SOCIALIZATION AFFECT RESEARCH QUALITY



Type of Doctorate
+
+
Practice-Oriented Socialization
Graduate Research Experience
-
+
+
Career Research Experience
-
+
+
Research Orientation
+
+
+
Research Quality

TABLE VI-15

ORIENTATION BY CAREER RESEARCH EXPERIENCE BY PRACTICE-ORIENTED SOCIALIZATION BY TYPE OF DOCTORATE

Orientation	Type of Doctorate																	
	Behavioral Sciences									Education								
	Amount of Practice-Oriented Socialization																	
	Less			Medium			More			Less			Medium			More		
	Career Research Experience																	
	Hi	Med	Lo	Hi	Med	Lo	Hi	Med	Lo	Hi	Med	Lo	Hi	Med	Lo	Hi	Med	Lo
Research	75	54	41	-	-	-	-	-	-	64	53	36	60	32	13	55	29	13
Teaching or Administration	15	33	46	-	-	-	-	-	-	18	33	56	36	52	73	33	50	64
Service or Other	10	14	13	-	-	-	-	-	-	18	13	8	5	16	15	12	21	23
	100%	101%	100%							100%	99%	100%	101%	100%	101%	100%	100%	100%
	(81)	(58)	(46)	(10)	(10)	(16)	0	(1)	(5)	(22)	(30)	(25)	(42)	(50)	(102)	(33)	(42)	(78)

All three variables--type of doctorate, less practice-oriented socialization, and career research experience--are independently related to greater research orientation (Table VI-15). Of the three, career research experience has the strongest effect on research orientation. Among behavioral science doctorates with less practice-oriented socialization, 75 per cent of those with considerable career research experience are research oriented compared to 41 per cent of those with little research experience. The relationship is similar and even stronger among education doctorates. Aside from showing the relatively greater importance of career research experience in affecting research orientation, Table VI-15 also reveals that research experience and less practice-oriented socialization compensate for each other in enhancing research orientation. That is, even if someone had more practice-oriented socialization, he might still become research oriented if he had a lot of research experience in his career. Similarly, if he did not have much research experience in his career, but he also did not have much practice-oriented socialization, then that increases his chance of being research oriented.

Based on the limited evidence presented in Table VI-15 it seems that type of doctorate also increases the likelihood of being research oriented, independently of less practice-oriented socialization and career research experience. In brief, then, several experiences increase the probability that authors will be research oriented, but the single most efficacious is career research experience. This means that whether an author received a doctorate in education or the behavioral sciences, and whether or not he had practice-oriented socialization, if he had a lot of research experience in his career, he is very likely to be research oriented. The additive nature of these experiences is illustrated throughout Table VI-15, but is seen most vividly when we compare authors at opposite extremes. Among behavioral science doctorates with less practice-oriented socialization and considerable career research experience, 75 per cent are research oriented, whereas among education doctorates with more practice-oriented socialization and little career research experience only 13 per cent are research oriented. When it is remembered that type of doctorate is related to practice-oriented socialization and to career research experience, these figures should indicate what is meant by a "developmental sequence" that fosters one type of orientation rather than another.

We have just seen that type of doctorate, less practice-oriented socialization, and more career research experience contribute independently to research orientation. Of the three, career research experience is most strongly related to research orientation. Since career research experience is strongly affected by graduate research experience, we wonder what the relative importance is for research orientation of research experience at two different stages of professional development. Authors become research oriented as a result of research experience at either stage of their development, although the more recent career research experience is more strongly related to research orientation than is graduate research experience (Table VI-16).

TABLE VI-16

ORIENTATION BY GRADUATE RESEARCH EXPERIENCE[a] BY CAREER RESEARCH EXPERIENCE

Orientation	Career Research Experience								
	High			Medium			Low		
	Graduate Research Experience								
	Both[b]	Ass'nt Only	Neither	Both	Ass'nt Only	Neither	Both	Ass'nt Only	Neither
Research	67%	65	54	55	52	25	47	19	20
Teaching or Administration	21	25	33	31	34	45	38	65	67
Service or Other	12	10	13	14	14	31	16	16	13
	100%	100%	100%	100%	100%	101%	101%	100%	100%
	(76)	(79)	(24)	(58)	(71)	(36)	(32)	(74)	(82)

[a] We examined this table while controlling for doctoral departmental prestige as well (because the variable was necessary for graduate research experience to enhance quality), but doctoral departmental prestige did not affect orientation so we did not include it in our presentation.

[b] Both assistant and in research center.

Among authors without graduate research experience, 54 per cent of those with considerable career research experience are research oriented compared to 20 per cent of those with little career research experience. Among authors with little career research experience, 47 per cent of those with both types of graduate research experience are research oriented compared to 20 per cent of those with neither type of graduate research experience. To some extent research experience at one time can compensate for the lack of it at another time, insofar as it affects research orientation. Not surprisingly, authors with both graduate research experience and career research experience are more likely to be research oriented than those with research experience at just one stage of their careers.

We know that type of doctorate affects graduate research experience (i.e., work as a research assistant and in a research center) and that graduate research experience increases the chance of career research experience. But we do not know if type of doctorate affects career research experience independently of graduate research experience. Behavioral science doctorates are more likely than education doctorates to have graduate research experience, and people with graduate research experience are more likely to have career research experience (which in turn affects research orientation which affects quality, as we have just seen). But are behavioral science doctorates more likely than education doctorates to have career research experience, even when graduate research experience is held constant? Does type of doctorate continue to affect career activities in addition to its effect through graduate research experience?

Our data reveal that if education doctorates had both types of graduate research experience, they are just as likely as behavioral science doctorates to have considerable career research experience. This indicates that type of doctorate does not affect career research experience providing that one has both types of graduate research experience (Table VI-17). It seems that education doctorates need a stronger dose of research experience during graduate training than do behavioral science doctorates in order to attain the same amount of career research experience.

In sum, having an education doctorate seems to have adverse effects on research experience in two ways. First, education doctorates are less likely to have graduate research experience, which decreases the likelihood that they will have considerable career research experience. Second, even if education doctorates do have graduate research experience, they need more than behavioral science doctorates in order to do a substantial amount of research in their careers. Among those with both experiences, 52 per cent of behavioral science doctorates and 44 per cent of education doctorates have had considerable career research experience, whereas among those who were only research assistants, 44 per cent of behavioral science doctorates compared to only 25 per cent of education doctorates have had considerable career research experience.

Moving away from the research experience portion of the model (Figure VI-1) toward the left side, we can consider an interpretation offered earlier in this chapter. When we were trying to explain why

TABLE VI-17

CAREER RESEARCH EXPERIENCE BY GRADUATE RESEARCH
EXPERIENCE BY TYPE OF DOCTORATE

Amount of Career Research Experience	Type of Doctorate					
	Behavioral Sciences			Education		
	Type of Graduate Research Experience					
	Both	Ass'nt but Not in Bureau	Neither	Both	Ass'nt but Not in Bureau	Neither
HIGH	52%	44	20	44	25	17
MEDIUM	35	36	22	34	30	26
LOW	13	20	59	22	45	57
	100%	100%	101%	100%	100%	100%
	(60)	(86)	(41)	(96)	(125)	(93)

certain experience affects research adversely, we mentioned that because of their practice-oriented nature, school teaching, the Ed.D. degree, and professional education courses may lessen an author's research orientation. If this proposed linkage is true, then the relationship between less practice-oriented socialization and better research should disappear when research orientation is held constant.

As expected, when authors are research oriented, practice-oriented socialization does not affect the quality of their research (Table VI-18). Among the research oriented, 42 per cent of those with less practice-oriented socialization are doing good research, compared to 45 per cent of those with some such socialization. Thus, if authors become research oriented despite having had some practice-oriented socialization, they are just as likely to do good research as those with less practice-oriented socialization. But among those who are not research oriented, having less practice-oriented socialization helps the quality of their research. Among the teaching oriented, for example, 33 per cent of those with less practice-oriented socialization are doing good research, compared to 15 per cent of those with more such socialization. This suggests that people with less practice-oriented socialization may be more research oriented than their counterparts, even though teaching is their primary orientation. In brief, our findings indicate that having less practice-oriented socialization benefits quality through two processes. The most important one is the way it increases the likelihood that authors will become research oriented. When that happens, authors do much better research. Even

TABLE VI-18

RESEARCH QUALITY BY PRACTICE-ORIENTED SOCIALIZATION BY ORIENTATION

	Orientation								
	Research			Teaching or Administration			Service or Other		
	Amount of Practice-Oriented Socialization								
	Less	Medium	More	Less	Medium	More	Less	Medium	More
ABOVE AVERAGE contribution to theory	47% (73)	47 (36)	[29] (14)	32 (31)	32 (71)	12 (50)	19 (21)	21 (24)	5 (19)
ABOVE AVERAGE contribution to practice	34 (73)	42 (36)	[64] (14)	35 (31)	28 (71)	18 (50)	33 (21)	21 (24)	16 (19)
ABOVE AVERAGE use of research methods	46 (72)	47 (36)	[50] (14)	32 (31)	23 (71)	14 (50)	33 (21)	21 (24)	11 (19)

if they do not become research oriented, however, authors with less practice-oriented socialization do better research than those with more, although proportionately they never match the research oriented.

The preceding analysis suggests the importance of research orientation for better research. Earlier we noted that authors with career research experience do better research than those without such experience. Since research experience is strongly related to research orientation, the question arises whether each of these variables is independently related to better research.

Our analysis shows that research experience and research orientation are independently related to better research, but research orientation has a stronger relationship (Table VI-19). Somewhat surprisingly, this means that authors who are research oriented but have not had much research experience in their careers do better research than those with a lot of experience who are not research oriented; 41 per cent of the former compared to 29 per cent of the latter are doing good research. Research oriented authors must somehow acquire the skills necessary to do better research, even if they have not had much research experience in their careers. They might also tend to be younger and thus not have had time to acquire much career research experience.

These data enhance the importance of research orientation for better research. Research orientation may be regarded as the crucial outcome of role socialization. As such it may also be considered a summary indicator of an author's role socialization.

TABLE VI-19

RESEARCH QUALITY BY CAREER RESEARCH EXPERIENCE
BY ORIENTATION

Research Quality	Orientation								
	Research			Teaching or Administration			Service or Other		
	Career Research Experience								
	High	Medium	Low	High	Medium	Low	High	Medium	Low
ABOVE AVERAGE contribution to theory	51% (63)	38 (34)	44 (23)	[29][a] (17)	19 (36)	23 (73)	- (11)	- (16)	7 (28)
ABOVE AVERAGE contribution to practice	46 (63)	35 (34)	35 (23)	[35] (17)	28 (36)	22 (73)	- (11)	- (16)	29 (28)
ABOVE AVERAGE use of research methods	47 (62)	50 (34)	44 (23)	[24] (17)	17 (36)	23 (73)	- (11)	- (16)	11 (28)

[a] The numbers forming the base of percentages is so small that the results must be considered highly tentative, hence we have enclosed the percentages in brackets.

Treating orientation as a summary measure of socialization permits us to return to a question raised earlier in this chapter. Does type of doctorate affect research quality independently of role socialization (as a result, for example, of differences in theoretical development), or is the relationship between type of doctorate and quality explained entirely by differences in socialization? We can consider these possibilities in Table VI-20. Type of doctorate is related to contribution to theory even when research orientation is held constant, suggesting that substantive differences between the two fields may exist and affect research quality. Among the research oriented, 49 per cent of behavioral scientists compared to 38 per cent of education doctorates are contributing to theory. With respect to theory, then, it is plausible that substantive differences between the two types of doctorates may help explain why behavioral scientists do better research than education doctorates. With respect to contribution to practice and use of research methods, however, education doctorates are doing just as well as behavioral science doctorates, when research orientation is held constant. There is virtually no difference in the proportion of behavioral scientists and education doctorates that contribute to educational practice and use sound research methods (43 per cent of the former and 42 per cent of the latter). It seems that type of socialization (as indicated by orientation) does explain why behavioral scientists are contributing more to practice and using better research methods than education doctorates.

TABLE VI-20

RESEARCH QUALITY BY TYPE OF DOCTORATE BY ORIENTATION

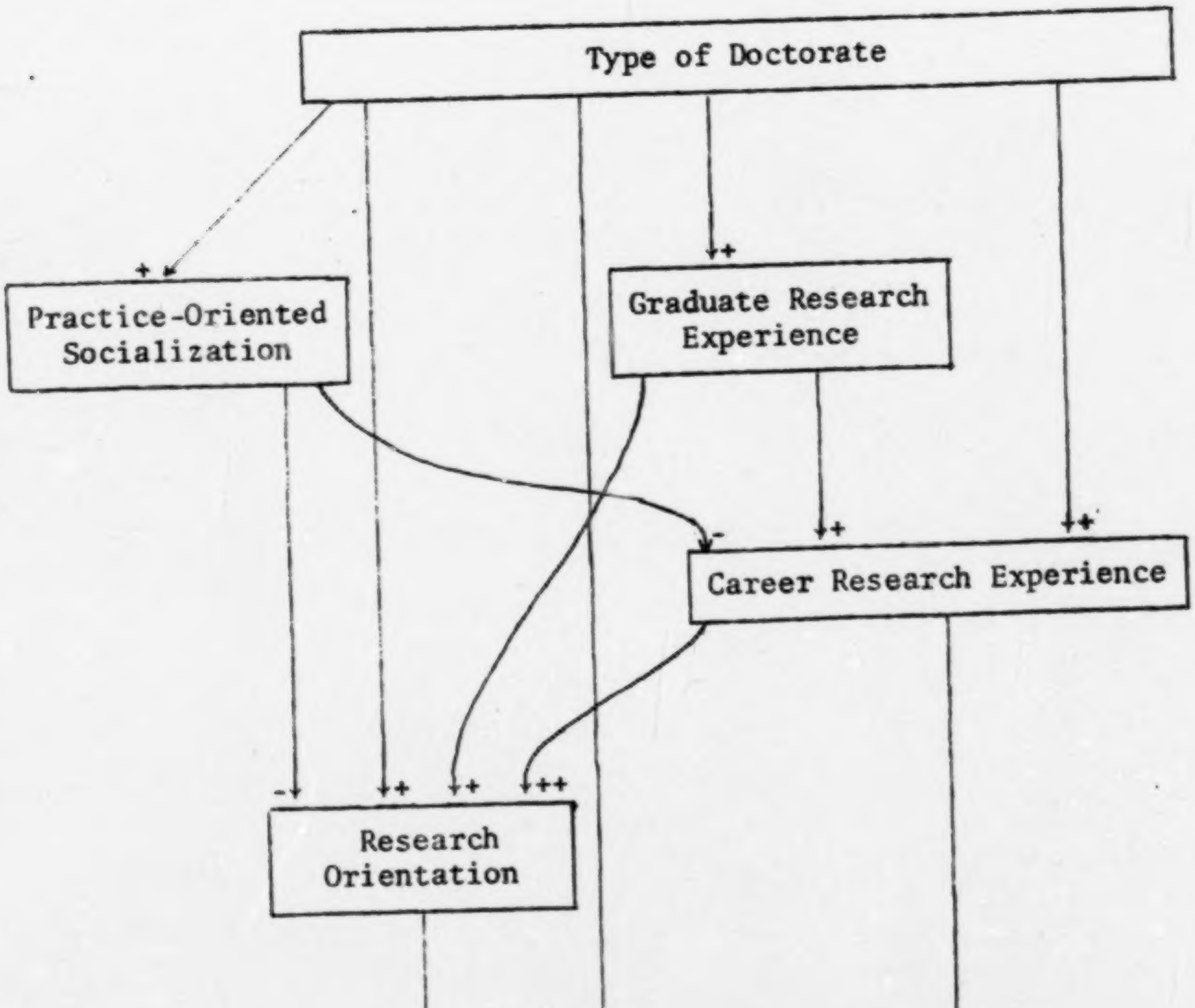
Research Quality	Orientation					
	Research		Teaching or Administration		Service or Other	
	Type of Doctorate					
	Behavioral Sciences	Education	Behavioral Sciences	Education	Behavioral Sciences	Education
ABOVE AVERAGE contribution to theory	49% (55)	38 (58)	28 (29)	25 (108)	[25] (12)	13 (32)
ABOVE AVERAGE contribution to practice	38 (55)	38 (58)	38 (29)	22 (108)	[33] (12)	19 (32)
ABOVE AVERAGE use of research methods	46 (55)	47 (57)	31 (29)	21 (108)	[33] (12)	22 (32)

In sum, in this chapter we have discussed five elements of role socialization and presented an hypothesized model of their relationships to each other and to quality. In the model we postulated that an author's socialization tended to direct him either toward the practice-oriented role of professional educator or toward the role of researcher, with major consequences for the quality of his research. Thus, education doctorates are more likely to have practice-oriented socialization, less likely to have graduate and career research experience and less likely to become research oriented, all of which has a cumulative negative effect on the quality of their research. In contrast, behavioral science doctorates are less likely to receive practice-oriented socialization, more likely to have graduate and career research experience and hence more likely to become research oriented, which has a positive cumulative effect on quality.

In general, the results of our analysis confirm the hypothesized model (in Figure VI-1) of how socialization affects quality. Only a few minor modifications of the model are required. First we add an arrow from graduate research experience to research orientation (Figure VI-2) to indicate that it affects orientation independently of career research experience. We put two "plus" signs by the arrow between career research experience and research orientation, to indicate that when antecedent variables are held constant, career research experience has the strongest relationship to research orientation. Finally, we put two "plus" signs by the arrow between research orientation and quality, to indicate that orientation is more strongly related to quality than is career research experience or type of doctorate.

FIGURE VI-2

MODIFIED MODEL OF HOW ROLE SOCIALIZATION AFFECTS QUALITY


Type of Doctorate
+
Practice-Oriented Socialization
+
Graduate Research Experience
-
+
+
Career Research Experience
-
+
+
++
Research Orientation
+++
+*
+
Research Quality

*Only with respect to contribution to theory

CHAPTER VII

BACKGROUND AND QUALITY

Every profession differentially recruits people of different backgrounds. Traditionally educational research has been performed by upwardly mobile, middle aged, white, male Protestants.[1] There is evidence and opinion suggesting that sex, religion, SES, age and academic achievement may affect quality, but the evidence is not so definite that we can predict how these background characteristics are related to quality. In come cases we have more specific expectations than in others.

In view of recently increased interest in the role of women in society and the professions, we considered it important to compare male and female researchers within our sample. In a study of reading researchers, Wilder found females were less involved in research in their careers than males.[2] In other academic fields, Bernard found females generally published less than males.[3] Similarly, Astin found that female Ph.D.'s publish less than males.[4] In science, McClelland asserts that creative physical scientists are more likely to be male than female.[1] Rossi found that female scientists publish less than males, as did David.[2] Only Simon has found that female scientists are just as likely as males to publish.[3] In view of these rather one-sided findings, it would be plausible to expect males to be doing better research than females, were it not for the lack of relationship we observe in Chapter IX (Table IX-1) between quality and productivity. This uncertainty makes it all the more worthwhile to compare male and female researchers in terms of the quality of their work.

---

[1] See Bargar, National Register of Educational Researchers.

[2] Wilder, "The Reading Experts."

[3] Jessie Bernard, Academic Women (New York: Meridian Books, World Publishing Co., 1964).

[4] Helen S. Astin, The Woman Doctorate in America (New York: Russell Sage Foundation, 1969).

With respect to religion as well as sex, earlier work has been more tantalizing than definitive. Wilder found that Jews are more involved in research,[4] and therefore it is plausible to expect Jews to do better research than members of other religions. In addition, we might expect the areligious to do better research than the religious, as noted in Chapter II. The expectation that the areligious will do better research is consistent with McClelland's generalization that creative physical scientists have a background of radical Protestantism but are not themselves religious.[1]

---

[1] David C. McClelland, "On the Psychodynamics of Creative Physical Scientists" in Contemporary Approaches to Creative Thinking, ed. by H. E. Gruber, G. Terrall, and M. Wertheimer (New York: Atherton Press, 1962), pp. 141-74.

[2] Alice S. Rossi, "Women in Science: Why So Few?" Science, May 28, 1965, pp. 1196-1202; idem, "Barriers to the Career Choice of Engineering, Medicine or Science Among American Women," in Women and the Scientific Professions, ed. by Jacquelyn A. Mattfeld and Carol G. Van Aken (Cambridge, Mass.: M.I.T. Press, 1965), pp. 51-127; Deborah David, "Career Patterns and Values: A Study of Men and Women in Science and Engineering" (unpublished Ph.D. dissertation, Columbia University, 1971).

[3] Rita Simon, Shirley M. Clark, and Kathleen Galway, "The Woman Ph.D.: A Recent Profile," Social Problems, XV, No. 2 (Fall, 1967), 221-36.

[4] Wilder, "The Reading Experts."

With regard to socio-economic status (SES), Hagstrom[2] noted the greater upward mobility of professors of education compared to behavioral science professors. In view of all we know about the relation of SES to educational advantages and academic achievement, we might expect higher SES authors to do better research. Buswell, however,[3] found that the SES of education doctorates was unrelated to their research productivity, although both productive and non-productive researchers were upwardly mobile. This suggests that SES may be unrelated to research quality. On the other hand, if people of lower SES have earned doctorates they may be especially motivated and hard-working and thus do better research. We will consider these possibilities in our analysis.

While age is a background characteristic of authors, it interests us primarily insofar as it reflects trends in educational research over time. In this vein, Bloom observed in 1966 that "there are some indications that the quality of educational research personnel is improving . . . ."[1] Persell found that among AERA members, younger researchers were more likely to be principal investigators on projects funded from outside sources than were older researchers.[2] Moreover, Barber suggests that many people in education were trained when research was less important than now, thus implying that younger researchers may be better trained.[3] In view of these data and opinions, we expect to find younger authors doing better research.

---

[1] McClelland, op. cit.

[2] Hagstrom, "Educational Researchers, Social Scientists, and School Professionals," prepared for a Colloquium on Improving the Social and Communication Mechanisms of Educational Research, sponsored by the American Educational Research Association, Washington, D.C., Nov. 21-22, 1968. (Mimeographed.)

[2] Buswell, et al., Training for Educational Research, p. 19.

After exploring the ascribed statuses of sex, religion, SES, and age in relation to quality, we consider an important achieved status, namely undergraduate performance of recruits. This factor has been offered as a major explanation of educational research quality. It is frequently stated that the best way to improve educational research is to attract more talented recruits. The better undergraduate performance of people entering graduate fields other than education is noted by Rossi, Davis, and McKinley,[4] and provides an alternative explanation of why behavioral science doctorates do better research than education doctorates. We will explore both of these possibilities and consider the relative importance of undergraduate performance compared to other explanations of quality.

---

[1] Bloom, "Twenty-five Years of Educational Research," p. 32.

[2] Caroline Hodges Persell, "AERA Membership Data" (New York: Bureau of Applied Social Research, Columbia University, 1970). (Xeroxed.)

[3] Bernard Barber, in Corwin and Seider, "Patterns of Educational Research."

[4] Peter H. Rossi, James A. Davis, and Richard J. McKinley, "Social Characteristics of 1961 College Graduates Entering the Field of Education" (paper presented at the annual program of the National Society of College Teachers of Education, February 16, 1962, Chicago, Illinois), cited in Sieber, Organization of Educational Research, p. 262.

In brief, in this chapter we will consider sex, religion, SES, age, and undergraduate performance in relation to quality. It is possible that background interacts with role socialization with consequences for quality, so we introduce socialization where relevant in our analysis. We conclude the chapter by adding what we learn about background characteristics to the model presented in Chapter VI.

Sex

Since no one has studied sex in relation to research quality, to our knowledge, we had no clear expectations with respect to our sample. Among educational researchers, females do better research than males on all three dimensions of quality--theory, practice, and research methods (Table VII-1). Thirty-nine per cent of females compared to 29 per cent of males are doing good research. Females are no more likely than males to be high undergraduate performers, so differences in the quality of their research are not explained by differences in undergraduate performance. It is possible that differences in quality are due to the greater tendency of females to earn doctorates in the behavioral sciences, since behavioral scientists do better research than education doctorates. Although there are too few examples to be conclusive, the data suggest that females do better research _only_ when they have received their doctorates in the _behavioral sciences_ (Table VII-2). Among behavioral science doctorates, 64 per cent of females do better research compared to 36 per cent of males.

TABLE VII-1

RESEARCH QUALITY BY SEX

Research Quality	Sex	
	Female	Male
Contribution to theory:		
Above average	37	29
Average	29	27
Below average	34	44
	100%	100%
	(49)	(301)
Contribution to practice:		
Above average	41	29
Average	29	36
Below average	30	35
	100%	100%
	(49)	(301)
Use of research methods:		
Above average	39	30
Average	25	34
Below average	37	37
	100%	100%
	(49)	(300)

TABLE VII-2

RESEARCH QUALITY BY SEX
BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate			
	Behavioral Sciences		Education	
	Sex			
	Female	Male	Female	Male
ABOVE AVERAGE contribution to theory	[54%] (13)	38 (84)	30 (23)	26 (178)
ABOVE AVERAGE contribution to practice	[69] (13)	33 (84)	26 (23)	26 (178)
ABOVE AVERAGE use of research methods	[69] (13)	36 (84)	22 (23)	30 (177)

The tendency of females to do better research is not explained by their greater career research experience either, since males and females have equal amounts of career research experience. How then can we explain these findings? Perhaps the greater self-selection of academic careers by females and the more formidable odds they face in pursuing academic careers[1] result in female authors who are somewhat differently oriented toward research than males. To analyze this possibility, we asked authors, "How did the topic of the paper mentioned above come to your attention?" Females are more likely than males to indicate that the most important source of their research topic was their "own field observation of a concrete problem in education"; 22 per cent of females compared to 14 per cent of males cited this as the most important way the research topic came to their attention. But when asked, "What was it about this topic which attracted you?," females are less likely than males to indicate that it was because it was "important for solving a concrete problem in education"; 25 per cent of females compared to 31 per cent of males mention this reason. Furthermore, females are more likely to indicate they selected a research topic because of its importance for developing or testing an educational or behavioral science theory; 27 per cent of females compared to 20 per cent of males cite this as the most important attraction of the topic. (See Table 1, 2 in Appendix VII-1.) So, even though females are more likely than males to have topics come to their attention from their own field observation of a concrete problem in education, they are also more likely to seize upon the research topic because of its importance for developing or testing a theory. This finding suggests that females may bring a more intellectual orientation to their research.

---

[1] See, for example, Rossi, "Women in Science" and "Barriers"; Pamela Roby, "Structural and Internalized Barriers to Women in Higher Education" (Department of Sociology and Center for Manpower Policy Studies, George Washington University, 1970). (Mimeographed.) Cynthia Epstein, Woman's Place: Options and Limits in Professional Careers (Berkeley: University of California Press, 1971).

Further evidence of a greater intellectual orientation among females can be found in what motivates them to do research more generally. The questionnaire item used to explore this possibility read:

> Of course, many factors motivate individuals to do research. Regarding your usual research motives, how would you distribute 10 points among the following (so the total equals 10)?

Seventy-five per cent of females compared to 65 per cent of males indicated that they are motivated to do research by "curiosity about substantive questions in a behavioral science field." Thus, it may well be the greater intellectual orientation of females that explains the interesting findings reported above.

When we test this interpretation by holding constant intellectual orientation, we find that males are just as likely as females to do better research when they are intellectually oriented (as measured by being motivated by curiosity about substantive questions in a behavioral science field). Among the intellectually oriented, 41 per cent of females and 40 per cent of males do better research (Table VII-3). In short, the generally greater intellectual orientation of female authors in our sample seems to explain why they do better research than males.

TABLE VII-3

RESEARCH QUALITY BY SEX BY INTELLECTUAL ORIENTATION

Research Quality	Intellectual Orientation					
	High		Medium		Low	
	Sex					
	Female	Male	Female	Male	Female	Male
ABOVE AVERAGE contribution to theory	[50%] (14)	46 (77)	38 (21)	26 (116)	[25] (12)	24 (102)
ABOVE AVERAGE contribution to practice	[29] (14)	36 (77)	43 (21)	29 (116)	[42] (12)	24 (102)
ABOVE AVERAGE use of research methods	[43] (14)	38 (76)	43 (21)	29 (116)	[25] (12)	25 (102)

Religion

As already noted, the literature shows that religion is related to research experience and orientation. We consider this finding from earlier research a very curious one. We wanted to know if it was corroborated by the data in our study, and if so, if we could offer an explanation of it. We measured religion in the general background section of the questionnaire (see Question 4.4 in the Questionnaire in Appendix II-3). Authors were asked, "Birthdate," "Sex," "Race," "Religion (Optional)." Because of the way the questionnaire item was phrased, we are unable to say whether religion reflects religious background or current affiliation. However, since the importance of religion for quality is probably explained by religion being an indicator of an underlying cultural or cognitive orientation rather than by something inherent in the particular value system of a religion, the religious classification provided by this questionnaire item is adequate for the purpose of our analysis. From now on when we use the term religion or religious classification, it should be thought of as an indicator of an underlying phenomenon, which we hope to specify through our analysis.

With respect to the relationship between religion and quality, our data are consistent with Wilder's findings. Religious classification is related to better research: 40 per cent of Jews, 37 per cent of the areligious and 28 per cent of Protestants are doing good research (Table VII-4). In an effort to determine what underlying phenomenon is associated with religious classification we systematically controlled for each background and role socialization characteristic that was related to religion.

TABLE VII-4

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION

	Religious Classification				
	Jewish	Areligious	Prot.	Cath.	Other
Contribution to theory:					
Above average	41%	39	30	27	20
Average	22	19	26	32	35
Below average	36	42	44	40	45
	99% (36)	100% (64)	100% (152)	99% (47)	100% (29)
Contribution to practice:					
Above average	37	36	27	34	27
Average	33	28	40	32	41
Below average	30	36	33	34	31
	100% (36)	100% (64)	100% (152)	100% (47)	99% (29)
Use of research methods:					
Above average	41	37	26	30	28
Average	25	36	33	34	38
Below average	33	27	42	36	34
	100% (36)	100% (64)	101% (29)	100% (151)	100% (47)

Since it is possible that the greater youth of Jewish researchers explains their better research (as we see later in this chapter, younger authors generally do better research) we controlled first for age. Even among younger researchers, however, Jews and the areligious do better research than Protestants (Table VII-5): 54 per cent of Jews, 41 per cent of the areligious and 26 per cent of Protestants do better research. But instead of explaining the relationship between religion and quality, youth increases that association.

If youth does not explain the relation between religious classification and quality, perhaps undergraduate performance (this concept is discussed later in this chapter) does so since it varies by religious classification. Undergraduate performance, however, does not explain why Jews and the areligious do better research (Table VII-6). Rather, high undergraduate performance and being Jewish or areligious are "conditionally interdependent" in their relation to quality, that is, one alone has only a small relationship to better research, but the two together result in much better research. Among high performers, 43 per cent of the areligious compared to 23 per cent of Protestants are doing better research. It seems that high undergraduate performance does not enhance the research quality of Protestants. Jews and the areligious, however, do better research than Protestants only when they were high undergraduate performers.[1]

---

[1] It is precisely this type of contingent relationship that is lost in such modes of analysis as multiple regression or path analysis. There are too few Catholics and other religions to draw any conclusions about them.

TABLE VII-5

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION BY AGE

Research Quality	Age											
	Under 40				40 - 49				50 or More			
	Religious Classification[a]											
	Jew.	Arel.	Prot.	Cath.	Jew.	Arel.	Prot.	Cath.	Jew.	Arel.	Prot.	Cath.
ABOVE AVERAGE contribution to theory	[61%] (18)	42 (31)	29 (70)	30 (34)	- (14)	39 (26)	33 (49)	- (9)	- (4)	- (8)	27 (33)	- (4)
ABOVE AVERAGE contribution to practice	[44] (18)	35 (31)	23 (70)	38 (34)	- (14)	35 (26)	33 (49)	- (9)	- (4)	- (8)	27 (33)	- (4)
ABOVE AVERAGE use of research methods	[56] (18)	45 (31)	26 (69)	32 (34)	- (14)	35 (26)	33 (49)	- (9)	- (4)	- (8)	15 (33)	- (4)

[a] 29 people who classified themselves as being some other religion were eliminated from this table because there were too few to subdivide by age.

TABLE VII-6

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION BY UNDERGRADUATE PERFORMANCE

Research Quality	Undergraduate Performance									
	High					Average				
	Religious Classification									
	Jewish	Arelig.	Prot.	Cath.	Other	Jewish	Arelig.	Prot.	Cath.	Other
ABOVE AVERAGE contribution to theory	[57%] (14)	38 (21)	23 (39)	- (6)	- (3)	32 (22)	39 (44)	32 (113)	22 (41)	19 (26)
ABOVE AVERAGE contribution to practice	[36] (14)	38 (21)	18 (39)	- (6)	- (3)	36 (22)	34 (44)	30 (113)	34 (41)	27 (26)
ABOVE AVERAGE use of research methods	[50] (14)	52 (21)	28 (39)	- (6)	- (3)	36 (22)	32 (44)	25 (112)	29 (41)	27 (26)

As discussed in the previous chapter, role socialization is a complex phenomenon consisting of type of doctorate, practice-oriented socialization, graduate research experience (which we will not consider here because it is related to better research through research orientation and career research experience, as we saw in the previous chapter), career research experience, and research orientation. All of these vary by religion and therefore the possibility arises that role socialization explains the relationship between religion and research quality.

Jews and the areligious are more likely to have behavioral science doctorates than Protestants, which might explain why they do better research. However, even among behavioral science doctorates, Jews and the areligious do better research than Protestants (Table VII-7). Having a behavioral science doctorate increases the difference between religious groups rather than lessening it. Among behavioral scientists, 51 per cent of the areligious compared to 29 per cent of Protestants are doing better research.

Jewish and areligious authors are less likely to have had practice-oriented socialization than are Protestants, which might explain their tendency to do better research. Even among those with less practice-oriented socialization, however, Jews and the areligious do better research than Protestants: 50 per cent of Jews and 49 per cent of the areligious do better research, compared to 30 per cent of Protestants (Table VII-8). Therefore it is not differences in practice-oriented socialization that explain the variation in quality associated with religious classification.

TABLE VII-7

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate							
	Behavioral Sciences				Education			
	Religious Classification							
	Jewish	Arelig.	Prot.	Catholic	Jewish	Arelig.	Prot.	Catholic
ABOVE AVERAGE contribution to theory	[57%] (14)	57 (23)	26 (39)	- (9)	[31] (13)	27 (33)	31 (93)	25 (32)
ABOVE AVERAGE contribution to practice	[50] (14)	52 (23)	28 (39)	- (9)	[15] (13)	30 (33)	23 (93)	31 (32)
ABOVE AVERAGE use of research methods	[57] (14)	44 (23)	33 (39)	- (9)	[39] (13)	36 (33)	24 (92)	28 (32)

TABLE VII-8

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION BY PRACTICE-ORIENTED SOCIALIZATION

Research Quality	Amount of Practice-Oriented Socialization											
	Low				Medium				High			
	Religious Classification											
	Jew.	Arel.	Prot.	Cath.	Jew.	Arel.	Prot.	Cath.	Jew.	Arel.	Prot.	Cath.
ABOVE AVERAGE contribution to theory	50% (22)	52 (33)	33 (43)	[39] (13)	- (7)	[28] (18)	41 (61)	33 (21)	- (6)	[25] (12)	13 (46)	[8] (12)
ABOVE AVERAGE contribution to practice	50 (22)	42 (33)	23 (43)	[46] (13)	- (7)	[22] (18)	34 (61)	33 (21)	- (6)	[17] (12)	10 (46)	[25] (12)
ABOVE AVERAGE use of research methods	50 (22)	52 (33)	33 (42)	[39] (13)	- (7)	[28] (18)	30 (61)	24 (21)	- (6)	[25] (12)	15 (46)	[33] (12)

Jews and the areligious also differ from Protestants in having more career research experience. This might explain why they do better work, but it does not. However, Jews and the areligious usually do not do better research than Protestants unless they have considerable career research experience (Table VII-9). When they have less career research experience, the areligious do no better research than Protestants. Among those with considerable career research experience, 53 per cent of the areligious do good research compared to 27 per cent of Protestants; whereas among those with less research experience, 27 per cent of the areligious compared to 25 per cent of Protestants are doing good work. That authors with more research experience do better work is not surprising. But it is unexpected that this relationship is contingent upon religious classification. Research experience leads to better research only among the areligious (and perhaps among Jews although there are too few examples to be sure), but not among Protestants. Protestants with considerable career research experience do no better work than those with less experience. Among Protestants, 27 per cent of the very experienced do good research, compared to 28 per cent of the less experienced. Not only does religion have a positive relationship to quality, but it seems to be a necessary condition for a relationship between research experience and better research.

Why is it that Protestants with considerable career research experience are not doing better work? One explanation may reside in their orientation. Protestants are less likely than Jews and the areligious to be research oriented, and this may be why they are less likely to do good research.

TABLE VII-9

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION
BY CAREER RESEARCH EXPERIENCE

Research Quality	Extent of Career Research Experience											
	High				Medium				Low			
	Religious Classification											
	Jew.	Arel.	Cath.	Prot.	Jew.	Arel.	Cath.	Prot.	Jew.	Arel.	Cath.	Prot.
ABOVE AVERAGE contribution to theory	[80] (10)	57 (23)	[46] (11)	29 (35)	[50] (12)	24 (21)	[46] (11)	24 (29)	[0] (10)	- (12)	15 (20)	30 (61)
ABOVE AVERAGE contribution to practice	[70] (10)	57 (23)	[46] (11)	23 (35)	[42] (12)	19 (21)	[55] (11)	31 (29)	[10] (10)	- (12)	25 (20)	28 (61)
ABOVE AVERAGE use of research methods	[40] (10)	44 (23)	[55] (11)	29 (34)	[58] (12)	38 (21)	[46] (11)	21 (29)	[10] (10)	- (12)	15 (20)	26 (61)

Research orientation does not explain why Jews and the areligious do better research, but it too is a necessary condition for that relationship to occur (Table VII-10). In other words, Jews and the areligious do better research than Protestants only when they are research oriented. Among the research oriented, 55 per cent of the areligious do good work compared to 28 per cent of Protestants. When Jews and the areligious are oriented toward teaching or administration, however, they are no more likely than Protestants to do better research; among the teaching oriented, 20 per cent of the areligious do good research compared to 27 per cent of Protestants.

Very strikingly, the strong relationship between research orientation and research quality is contingent upon religious classification, just as was the relationship between research experience and quality. Being research oriented is related to better work only among the areligious, Jews, and Catholics, but not among Protestants. Among Protestants, 28 per cent of the research oriented are doing good work compared to 27 per cent of the teaching oriented.

None of the aspects of role socialization we have considered explain why Jews and the areligious do better research than Protestants. We have learned that they do better research only when they are high academic achievers, have considerable research experience in their careers and when they are research oriented. We have also seen that Jews and the areligious do even better research when they have behavioral science doctorates and less practice-oriented experience during their professional socialization. Perhaps the most significant

TABLE VII-10

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION BY ORIENTATION

Research Quality	Orientation											
	Research				Teaching or Admin.				Service or Other			
	Religious Classification											
	Jew.	Arel.	Cath.	Prot.	Jew.	Arel.	Cath.	Prot.	Jew.	Arel.	Cath.	Prot.
ABOVE AVERAGE contribution to theory	[77] (13)	61% (28)	42 (19)	30 (43)	31 (16)	17 (23)	21 (19)	31 (81)	- (7)	- (12)	- (9)	- (21)
ABOVE AVERAGE contribution to practice	[69] (13)	39 (28)	42 (19)	26 (43)	19 (16)	26 (23)	37 (19)	27 (81)	- (7)	- (12)	- (9)	25 (24)
ABOVE AVERAGE use of research methods	[69] (13)	64 (28)	37 (19)	29 (42)	38 (16)	17 (23)	26 (19)	22 (81)	- (7)	- (12)	- (9)	33 (24)

finding in this section is the discovery that high academic achievement, more research experience, and research orientation are related to better research only among Jews and the areligious but not among Protestants. That is, these relationships are contingent upon religious classification. This mysterious finding warrants further investigation. As noted earlier, religious classification is probably an indicator of differences in cultural or cognitive orientations. Friedman's recent study of college work-study students offers evidence that religion is an indicator of some underlying orientation. In that study, Protestant students were more likely than Jewish or areligious students to indicate that the most important purpose of college is vocational (60 per cent of Protestant students stated this, compared to 47 per cent of Jewish students and 40 per cent of the areligious). On the other hand, Jewish and areligious students were more likely to consider the most important purpose of college as being intellectual (36 per cent of Jewish and areligious students compared to 26 per cent of Protestant students agreed to that).[1] It is important to note that these differences in opinion about the most important purpose of college occur even though areligious and Jewish students are more likely than Protestants to indicate "it will be hard to get a job" when they graduate. Thirty-six per cent of the areligious and 27 per cent of Jewish students feel this way, compared to only 14 per cent of Protestants. Thus, in some ways the areligious and Jewish students should be more likely than Protestants to consider the vocational purpose of college as most important. That they do not, even in the face of perceived practical difficulties, suggests something about the strength of their intellectual orientation.

---

[1] Nathalie Friedman, unpublished data from a comprehensive study of the U.S. Office of Education Educational Opportunity Grant Program. Prepared at the Bureau of Applied Social Research, Columbia University, New York. All of these respondents must qualify for financial aid in order to be named work-study job recipients. Therefore, all are of lower economic status and we may assume that social class differences do not account for the differences betwеen areligious groups noted above.

Friedman's data suggest that the areligious and Jews may be more intellectually oriented, while Protestants are more vocationally or pragmatically oriented. We can test this interpretation with data on how authors in this population selected their research topics and with information on what motivated them to do research. If this interpretation is correct, we would expect Protestants to be more likely to indicate pragmatic motivations and reasons for selecting a particular research topic, while Jews and the areligious would have more intellectual motivations and reasons for topic selection. We will look first at topic selection to see if this process varies by religious classification.

There are several important differences in topic selection between religious groups, and they are in the anticipated direction. Regarding the source of their research topic, we asked authors, "How did the topic of the paper mentioned above come to your attention?" Jews and the areligious are more likely than Protestants to indicate that the most important source of their topic was their "own reflection, i.e., not based on previous research or others' ideas"; 21 per cent of Jews and 18 per cent of the areligious compared to 13 per cent of Protestants indicate this was the most important source of their topic (see Table 3, Appendix VII-1).[1] Perhaps Jews and the areligious are more likely than Protestants to derive their research topic from an intellectual source.

We also asked authors, "What was it about this topic which attracted you?" The biggest difference in attraction of research topic among different religions occurs among those who liked the topic because it was "important for developing or testing a behavioral science theory." Twenty-four per cent of Jews and 20 per cent of the areligious indicate this reason, compared to only 8 per cent of Protes- (see Table 4 in Appendix VII-1). Again, Jews and the areligious are more likely than Protestants to cite an intellectual reason for selecting their research topic.

In addition, the areligious and Jews are *less* likely than Protestants or other religions to be attracted to a research topic because of its "importance for solving a concrete problem in education," which indicates a pragmatic orientation. Thirty-three per cent of Protestants compared to 26 per cent of Jews and 22 per cent of the areligious indicated this as the most important reason they were attracted to the topic (Table 4, Appendix VII-1). From these findings we conclude that Jews and the areligious are more likely to select their research topics for intellectual reasons, while Protestants are more likely to select research topics for pragmatic reasons.

---

[1] Although the percentage differences between religious groups with respect to topic selection may seem small, they are the largest variations in topic selection observed in the study. Most of the key variables in this study (for example, type of doctorate and research orientation) were not related to topic selection.

Furthermore, religious classification is also an indicator of more general motivations for doing research. We tapped motivations with the following questionnaire item:

> Of course, many factors motivate individuals to do research. Regarding your usual research motives, how would you distribute 10 points among the following (so the total equals 10)?

Seventy-eight per cent of Jews and 73 per cent of the areligious compared to 63 per cent of Protestants gave one or more points to being motivated by "curiosity about substantive questions in a behavioral science field," another indicator of their greater intellectual orientation (see Appendix VII-2).

These findings make it quite evident that religion is an indicator of what we can call intellectual orientation. The question remains--is it this intellectual orientation that explains why religious classification is related to quality? To test this interpretation, we examine the relationship between religion and research quality, while controlling for intellectual orientation (Table VII-11). And, indeed we find that intellectual orientation does reduce the relationship between religion and contribution to theory. Protestants are almost as likely as Jews or the areligious to contribute to theory provided they are very intellectually oriented. Differences between religious groups remain with respect to practice and methods, however, even when intellectual orientation is equally great (when the ratings of the two dimensions are averaged): 43 per cent of the areligious and

TABLE VII-11

RESEARCH QUALITY BY RELIGIOUS CLASSIFICATION BY INTELLECTUAL ORIENTATION

Research Quality	Intellectual Orientation											
	High				Medium				Low			
	Religious Classification											
	Jewish	Arelig.	Protes.	Cath.	Jewish	Arelig.	Protes.	Cath.	Jewish	Arelig.	Protes.	Cath.
ABOVE AVERAGE contribution to theory	[54%] (13)	48 (29)	47 (30)	- (7)	[47] (15)	[44] (18)	21 (63)	[29] (17)	- (8)	[17] (18)	33 (55)	[16] (19)
ABOVE AVERAGE contribution to practice	[54] (13)	38 (29)	30 (30)	- (7)	[27] (15)	[50] (18)	22 (63)	[47] (17)	- (8)	[17] (18)	29 (55)	[26] (19)
ABOVE AVERAGE use of research methods	[39] (13)	48 (29)	28 (29)	- (7)	[53] (15)	[33] (18)	22 (63)	[53] (17)	- (8)	[28] (18)	29 (55)	[11] (19)

47 per cent of Jews compared to 29 per cent of Protestants are contributing to educational practice and using good research methods.[1] Among the less intellectually oriented, Protestants are doing better research than the areligious; 30 per cent of Protestants compared to 21 per cent of the areligious are doing good research.[2]

In sum, the relation of religion to quality is partially explained by intellectual orientation. Furthermore, Jews and the areligious do better research only when they are intellectually oriented. If we return to the beginning of this section, and substitute the term intellectually oriented for Jews and the areligious, many of the findings in this section become more understandable. We can see now why the relationship between religious classification (which really seems to be an indicator of intellectual orientation) and research quality is contingent upon high academic achievement. It is not enough to be intellectually oriented, as are the Jews and the areligious, academic achievement is needed as well.

Further, we can comprehend why the relation between religion and quality is intensified by having a behavioral science doctorate. As noted in Chapter VI, obtaining a behavioral science doctorate is related to less practice-oriented socialization, to greater graduate and career research experience, and to greater research orientation.

---

[1] The remaining relationship between religious classification and quality might be due to the more cosmopolitan background of Jews compared to Protestants. We did not have the necessary data to test this hypothesis, however.

[2] There are too few Jews who are not intellectually oriented to make statements about them.

Moreover, we found evidence to support the idea that the behavioral sciences have a higher level of theoretical development than does education. Therefore, it is not surprising that the combination of religious classification and a behavioral science doctorate are related to even better research. Finally, we can understand why the relation of religion (i.e., intellectual orientation) is contingent upon career research experience. It is not enough to be intellectually oriented, one must have research experience if intellectual orientation is to enhance quality.

Socio-Economic Status

We measured the socio-economic status (SES) of authors in this sample by the educational level of their parents. Bargar did not collect this information in his study, but Buswell did. He found that only 13 per cent of the fathers and 8.7 per cent of the mothers of the researchers he studied were college graduates, and he concluded that the researchers were a very upwardly mobile population.[1] Similarly, among Wilder's reading researchers, the fathers of only 17 per cent were college graduates. As Hagstrom has pointed out,[2] only 28 per cent of university professors in educational fields were sons of fathers who had attended college, compared to 39 per cent of the social scientists in U.S. universities.[3] Hagstrom does not draw out the implications of this finding. He notes that those in educational fields are more likely to have made their career choice later in life, to have pursued a career with more changes in it, to have a characteristic ideology with a service emphasis, and to be less committed to a reference group of disciplinary colleagues, although he does not discuss the relationship of these characteristics to upward mobility.[1] Nor does he suggest how the greater upward mobility of people in education might affect the quality of their research.

---

[1] Buswell, et al., Training for Educational Research.

[2] Hagstrom, "Educational Researchers."

[3] Ibid.

Among authors in our sample, 29 per cent had fathers who attended or graduated from college. This is consistent with Hagstrom's finding since 38 per cent of this population are in the behavioral sciences. As noted at the beginning of this chapter, there is evidence for contradictory expectations about the relationship between SES and quality. The socio-economic status of the researcher's family has no striking relationship to research quality (Table VII-12). Twenty-six per cent of those whose fathers were elementary school graduates are doing good research compared to 21 per cent of those whose fathers attended graduate school. Authors whose fathers attended high school or college are doing slightly better research, but the differences are small and not worth analyzing in detail. Therefore we will not consider parental SES further in our analysis.

---

[1] Ibid.

TABLE VII-12

RESEARCH QUALITY BY SES (EDUCATION OF FATHER)

Research Quality	Education of Father			
	Elementary or Less	High School or H.S. Grad.	College or College Grad.	Some Graduate School or Grad. Degree
Contribution to theory:				
Above average	28%	32	35	19
Average	16	28	31	32
Below average	56	41	33	48
	100%	101%	99%	99%
	(64)	(148)	(102)	(31)
Contribution to practice:				
Above average	30	32	31	26
Average	30	35	40	29
Below average	41	32	28	45
	100%	99%	99%	100%
	(64)	(148)	(102)	(31)
Use of research methods:				
Above average	20	35	36	19
Average	28	30	38	32
Below average	52	35	26	48
	100%	100%	100%	99%
	(64)	(147)	(102)	(31)

Age

We are interested in age as an historical phenomenon, in terms of what it portends for educational research. As we noted in the beginning of this chapter, there is some evidence for expecting younger authors to do better research. In view of this it is not surprising that age has a negative linear relationship to quality. Among researchers under thirty, 36 per cent are doing good research, compared to only 22 per cent of researchers fifty or older (Table VII-13). Younger authors are doing better research than older ones primarily when they received doctorates in education rather than in the behavioral sciences (Table VII-14). Why might this be the case? Younger authors with both types of doctorates are more likely to be research oriented than are older ones (Table VII-15). When we examine the relationship between youth and quality, while holding constant research orientation, we see that the research orientation of younger authors explains their tendency to do better research, except with respect to theory (Table VII-16). Among the research oriented, 52 per cent of younger authors compared to 38 per cent of older ones are contributing to theory, but there is no difference according to age with respect to practice or use of research methods. These findings suggest that the relation between youth and quality is explained by the greater research orientation of younger authors, except with respect to theory. Among teaching oriented authors, however, the younger ones do considerably better research than older ones. This suggests that among younger authors, the teaching orientation may not be as detrimental for the quality of their research as it is among older authors.

TABLE VII-13

RESEARCH QUALITY BY AGE

Research Quality	Age			
	Under 30	30-39	40-49	50 or More
Contribution to theory:				
Above average	40%	34	28	22
Average	20	31	25	24
Below average	40	34	47	54
	100%	99%	100%	100%
	(25)	(151)	(121)	(54)
Contribution to practice:				
Above average	32	31	31	28
Average	28	38	34	32
Below average	40	31	35	41
	100%	100%	100%	101%
	(25)	(151)	(121)	(54)
Use of research methods:				
Above average	36	34	31	17
Average	36	33	31	35
Below average	28	33	38	48
	100%	100%	100%	100%
	(25)	(150)	(121)	(54)

TABLE VII-14

RESEARCH QUALITY BY AGE BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate					
	Behavioral Sciences			Education		
	Age					
	Under 40	40 - 49	50 or More	Under 40	40 - 49	50 or More
ABOVE AVERAGE contribution to theory	40% (53)	50 (28)	29 (17)	36 (95)	20 (75)	23 (31)
ABOVE AVERAGE contribution to practice	34 (53)	46 (28)	41 (17)	32 (95)	23 (75)	19 (31)
ABOVE AVERAGE use of research methods	43 (53)	36 (28)	35 (17)	33 (95)	32 (75)	10 (31)

TABLE VII-15

ORIENTATION BY AGE BY TYPE OF DOCTORATE

Orientation	Type of Doctorate							
	Behavioral Sciences				Education			
	Age							
	Under 30	30-39	40-49	50-70	Under 30	30-39	40-49	50-70
Research	56	57	53	41	50	34	24	24
Teaching or Administration	25	33	33	34	39	52	58	58
Service or Other	19	11	14	25	11	14	18	18
	100%	100%	100%	100%	100%	100%	100%	100%
	(16)	(120)	(88)	(32)	(18)	(217)	(194)	(72)

TABLE VII-16

RESEARCH QUALITY BY AGE BY ORIENTATION

Research Quality	Orientation								
	Research			Teaching or Administration			Service or Other		
	Age								
	< 37	38-42	> 43	< 37	38-42	> 43	< 37	38-42	> 43
ABOVE AVERAGE contribution to theory	52% (58)	41 (29)	38 (37)	34 (53)	26 (43)	17 (60)	18 (22)	6 (16)	19 (27)
ABOVE AVERAGE contribution to practice	41 (58)	38 (29)	41 (37)	38 (53)	21 (43)	18 (60)	9 (22)	13 (16)	41 (27)
ABOVE AVERAGE use of research methods	47 (57)	52 (29)	43 (37)	32 (53)	23 (43)	12 (60)	18 (22)	25 (16)	22 (27)

In sum, the tendency of younger education doctorates to do better research than older ones is largely explained by their greater research orientation. When research orientation is held constant, younger authors surpass older ones only with respect to contribution to theory. This suggests that the role socialization of younger authors somehow facilitates their contributing to theory with their research.

Undergraduate Performance

As noted in the beginning of this chapter, a number of observers have urged improving the quality of recruits entering educational research. Despite the tremendous emphasis on the importance of recruiting more talented people into research, however, undergraduate performance[1] has only a small positive relationship to better theory and methods, and a negative relationship to practice (Table VII-17). This small relationship should be qualified by one consideration, however. Those concerned about raising the caliber of recruits entering educational research are speaking of all entrants to the field, whereas we are analyzing a very small proportion of those originally recruited. Many people entering educational research do not become active researchers like the ones in this sample. Therefore, it is possible

---

[1] Undergraduate performance was measured as follows: First, undergraduate school selectivity was measured by using Berelson's ranking of colleges and universities, Cass and Birnbaum's evaluation of the admissions selectivity of colleges and universities, and an index of school resources developed at the Bureau of Applied Social Research. Second, we asked authors if they had won any academic honors as an undergraduate. If an author earned one or more academic honors in a highly selective undergraduate college or university, he was classified as being a high performer. If not, as being an average performer.

TABLE VII-17

RESEARCH QUALITY BY UNDERGRADUATE PERFORMANCE

Research Quality	Undergraduate Performance	
	High	Average
Contribution to theory:		
Above average	35%	29
Average	30	27
Below average	35	44
	100%	100%
	(88)	(264)
Contribution to practice:		
Above average	27%	32%
Average	41	33
Below average	32	35
	100%	100%
	(88)	(264)
Use of research methods:		
Above average	38	29
Average	32	33
Below average	31	39
	101%	100%
	(88)	(263)

that at this level undergraduate performance may be a relatively minor variable in explaining quality, while at the point of entry it is a very important factor.

A further issue remains. Are behavioral scientists doing better research than education doctorates because they were better undergraduate performers? Many people have assumed this to be the case. When we examine the data, however, we see that better undergraduate performers and average performers are about equally likely to earn doctorates in education and in the behavioral sciences (Table VII-18). Among high performers, 64 per cent earn education doctorates, while among average performers 67 per cent earn education doctorates. Better undergraduate performers are just as likely to have the same type of role socialization, in other respects besides type of doctorate, as are average performers.

TABLE VII-18

TYPE OF DOCTORATE BY UNDERGRADUATE PERFORMANCE

Type of Doctorate	Undergraduate Performance	
	High	Average
Education	64%	67
Behavioral Sciences	36	34
	100%	101%
	(204)	(564)

In sum, while undergraduate performance has a small positive relationship to contribution to theory and use of research methods, it explains much less variation in quality than does either role socialization or intellectual orientation. Therefore, we will not consider it further in our analysis.

Summary

In this chapter we have seen that females, the areligious, Jews, and younger authors are doing better research. The upwardly mobile do no better research than the socially stable, and high undergraduate performers do only slightly better research than average performers. The finding that females, the areligious, and Jews do better research is not explained by different patterns of role socialization, although those relationships are sometimes contingent upon or increased by socialization. Instead we discovered that intellectual orientation (as measured by being motivated to do research by substantive curiosity about questions in a behavioral science field) largely explains why females, the areligious, and Jews do better research. When intellectual orientation is held constant, males are just as likely as females and Protestants are almost as likely as Jews and the areligious to be doing better research. The major explanation of why younger authors do better research seems to be their greater research orientation. Among the research oriented, younger authors differ from older ones only with respect to their greater contribution to theory.

Our analysis so far suggests two key variables that affect research quality, namely intellectual orientation and research orientation. The question remains--are these two orientations independently related to better research? Our data indicate that both research orientation and intellectual orientation are independently related to better research (Table VII-19). Moreover, the presence of one increases the relationship between the other and quality. Authors who are research oriented are likely to do better research, especially if they are also intellectually oriented. Among the intellectually oriented, 48 per cent of the research oriented are doing good research compared to 25 per cent of the service oriented (a difference of 23 per cent) whereas among the less intellectually oriented 30 per cent of the research oriented do better research compared to 19 per cent of the service oriented (a difference of only 11 per cent). Similarly, being intellectually oriented is related to better research among all authors, but it helps quality even more among those who are research oriented. Among the research oriented, 48 per cent of the intellectually oriented are doing good research compared to 30 per cent of the less intellectually oriented (a difference of 18 per cent), whereas among the teaching oriented, 32 per cent of the intellectually oriented do good research compared to 24 per cent of the less intellectually oriented (a difference of only 8 per cent).

As we can see from the above figures, research orientation has a stronger relationship to quality than does intellectual orientation, when the two are considered together. These findings can be added to the flow chart from Chapter VI (see Figure VII-1). The positive interaction between intellectual orientation and research orientation in

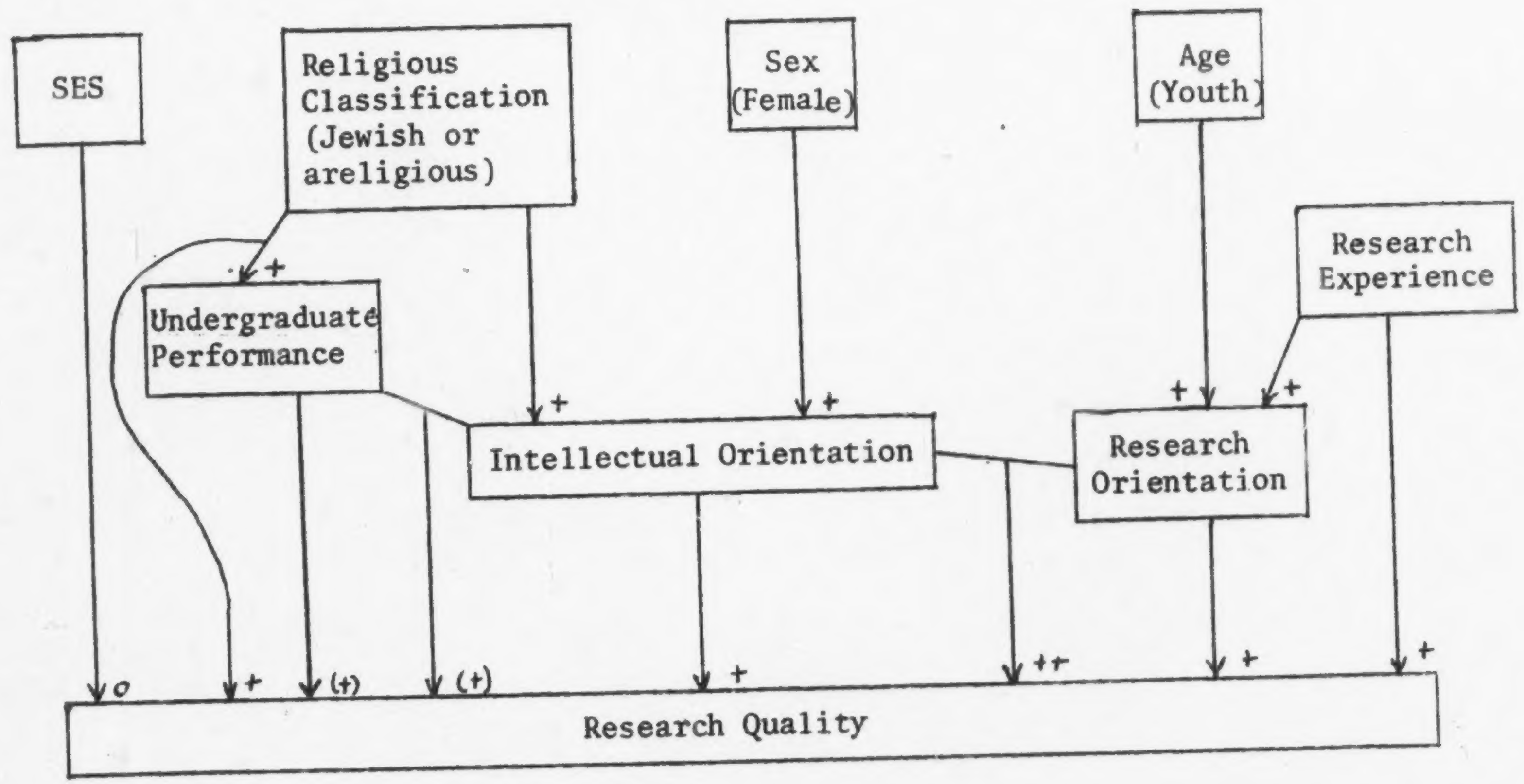
TABLE VII-19

RESEARCH QUALITY BY (ROLE) ORIENTATION
BY INTELLECTUAL ORIENTATION

Research Quality	Intellectual Orientation								
	High			Medium			Low		
	(Role) Orientation								
	Research	Teaching or Admin	Service or Other	Research	Teaching or Admin	Service or Other	Research	Teaching or Admin	Service or Other
ABOVE AVERAGE contribution to theory	61% (46)	31 (32)	[25] (12)	44 (48)	22 (60)	15 (27)	25 (28)	28 (58)	12 (25)
ABOVE AVERAGE contribution to practice	35 (46)	38 (32)	[25] (12)	48 (48)	23 (60)	22 (27)	32 (28)	22 (58)	24 (25)
ABOVE AVERAGE use of research methods	49 (45)	28 (32)	[25] (12)	52 (48)	20 (60)	22 (27)	32 (28)	22 (58)	20 (25)

FIGURE VII-1

INTERRELATION OF BACKGROUND, SOCIALIZATION, AND RESEARCH QUALITY


SES
Religious Classification (Jewish or areligious)
Sex (Female)
Age (Youth)
Research Experience
Undergraduate Performance
Intellectual Orientation
Research Orientation
Research Quality
+
+
+
+
+
o
+
(+)
(+)
+
++
+
+

relation to quality is indicated with the arrow drawn from the line connecting the two variables.

In the next chapter we direct our attention to the organizational settings in which the authors did their research, to see if settings increase our understanding of what affects research quality.

CHAPTER VIII

RESEARCH SETTINGS AND QUALITY

The evidence we have considered so far indicates that role socialization (as summarized by research orientation) and intellectual orientation are related to better research. An additional explanation of research quality has been offered by observers of educational research who emphasize the importance of the setting in which a researcher works. We will consider four settings which are frequently mentioned in relation to quality. These settings are: (1) university or other organizational affiliation, (2) school or department, (3) departmental prestige, and (4) affiliation with a university research center. We will briefly examine each of these settings.

A major issue in educational research is whether research should be conducted in universities, in specialized agencies which are independent of universities, or in some other setting. Some feel that the university is the best setting, because research should be conducted in a situation where close contact can be maintained with new conceptual and theoretical developments. On the other hand, others argue that independent, non-profit organizations offer unique opportunities for doing research. In a report to the Board of Directors of AIR, the following advantages were noted:

full-time research staff,
continuity of effort,
planning of desired objectives and priorities,
definitive studies,
availability of advanced technology,
advisory resources,
interdisciplinary team approach,
utilization of outstanding research talent,
flexibility of location,
evaluation of the research conducted there,
organizational grants for pilot research projects.[1]

These advantages would not be accepted by all observers of independent agencies, however. In a study of the organization for research in the social sciences made a few years ago by Young, the following conclusion was reached on the subject of independent agencies:

> There is no substantial evidence that research in the social sciences dissociated from academic life is more favorably situated for efficient operation by virtue of its isolation. Indeed, there seems to be nothing that an independent agency can do which cannot be done at least as well and possibly better in close association with an educational institution.[2]

In view of these differences of opinion, we have no definite expectations about the relationship between organizational affiliation and quality. At the same time, these opinions emphasize the importance of examining this issue.

Other observers[3] have suggested that the school or department in which research is conducted affects its quality. Schools of education are closely tied to the service needs of public schools, they argue, thereby creating a climate which impedes good research. The service orientation of schools of education means that research may be conducted under crisis-oriented circumstances, resulting in research that is under-conceptualized and thereby makes only a short-run contribution to educational problems. Also, the definition of research is affected by the climate of schools of education, so that school surveys and other service activities are often confused with research, as Sieber documented.[1] Further, the demands of school systems divert manpower from research into service activities. Finally, the close contact with educational systems may occasionally result in research undertaken to legitimize existing programs. These conditions probably affect research adversely.

---

[1] American Institutes for Research, Report to the Board of Directors, October 26, 1964. (Xeroxed.)

[2] Donald Young, "Organization for Research in the Social Sciences" (New York: Russell Sage Foundation, p. 20; mimeographed), cited in Sam D. Sieber, "Existing Organizational Patterns in Educational Research," in The Training and Nurture of Educational Researchers, p. 160.

[3] See for example, Corwin, "Patterns of Educational Research," Hagstrom, "Educational Researchers," and Sieber, "Organization of Educational Research."

It is generally assumed that better research will be produced in departments with higher prestige, because these departments can attract and select more talented faculty and students. In addition, Sieber's study[2] showed that schools of education with reputations for doing better research were more likely to have greater contact with the liberal arts, to have more facilitative arrangements for research, to emphasize research background over teaching experience in hiring of faculty, etc. Therefore, we expect that authors in more prestigious departments will have done better research.

---

[1] Sieber, Organization of Educational Research.

[2] Ibid.

Finally, students of research settings often remark that the best setting for research is a university research center. Such centers provide facilities and an intellectual climate for research, as well as offering valuable social supports for research. Moreover, in such a setting an author may find collaborators with skills and orientations which complement his own, thereby enhancing his research.

Sieber has described in greater detail these settings for educational research in the United States and the type of work which is conducted in them.[1] In this chapter we will examine the relationship between these settings and research quality to see if research settings provide a supplementary explanation of quality.

University or Other Organizational Setting

We assume that certain organizational settings, especially universities and specialized agencies,[2] provide more time, intellectual interaction, facilities, and possibly more funds for research than other settings. Therefore we expect that authors working in universities and in specialized agencies will do better research than authors working in other organizations. We asked authors,

> At the time of this research, what was your main institutional affiliation? University, college, school system, state department of education, private agency, or other (please specify). [See Question 1.10 in the Questionnaire, Appendix II-5.]

---

[1] Sam D. Sieber, "Institutional Settings for Educational Research in the U.S." in Educational Change in the U.S., ed. by Egon Guba (Bloomington, Ind.: Indiana University, 1967), prepared for UNESCO.

[2] These specialized agencies are non-governmental organizations such as Educational Testing Service or the American Institutes of Research.

As expected, authors who were working in a university or a specialized agency when they wrote their research paper produced[1] better research than those working in colleges, school systems, or other organizations (Table VIII-1). Forty-two per cent of those in specialized agencies, 34 per cent of those in universities, 21 per cent of those in colleges and 15 per cent of those in school systems did good research. University authors contributed more to theory and specialized agency authors used better research methods, and they both contributed more to educational practice than authors working in other organizations.

TABLE VIII-1

RESEARCH QUALITY BY ORGANIZATIONAL AFFILIATION

Research Quality	Organizational Affiliation				
	Specialized Agency	University	College	School System	Other
ABOVE AVERAGE contribution to theory	[33%] (15)	35 (264)	19 (36)	13 (30)	[7] (15)
ABOVE AVERAGE contribution to practice	[40] (15)	33 (264)	19 (36)	23 (30)	[20] (15)
ABOVE AVERAGE use of research methods	[53] (15)	33 (263)	25 (36)	10 (30)	[20] (15)

---

[1] The reader will note that in this chapter, unlike in other chapters, we present our results in the past tense, e.g., authors in particular settings did better research. We do this because we want to make it clear that we are considering the relationship between working in a given setting when a specific piece of research was done, and the rated quality of that particular research paper.

Although the number of authors who worked in specialized agencies is small, these data suggest that research does not need to be conducted in universities in order to contribute to theory. Authors in specialized agencies contributed virtually as much to theory as did university-based authors, and they surpassed university-based authors with respect to their contribution to practice and their use of research methods.

Authors in either universities or specialized agencies did much better research than authors affiliated with any other type of organization. The poor showing of authors in colleges, school systems, or other organizations suggests several factors that may be important for better research. First, we may infer that it is more difficult for authors in colleges to stay abreast of the theoretical, practical and methodological issues in educational research. Authors in school systems and other settings may be similarly isolated from the issues of research. Moreover, such authors may be required by their jobs to do research on particular problems, whether or not those questions contribute to theory or practice and whether or not methods exist (or are known to the author) for studying the problems. Finally, authors in all three of these types of settings may lack colleagues with whom to discuss their research problems.

It is possible that the above relationships between organizational affiliation and quality are really explained by the orientations of authors who worked in those settings. In the previous chapters we noted two orientations which are strongly related to quality, namely intellectual orientation and research orientation. Authors affiliated with all of the above types of organizations were equally likely to be intellectually oriented (see Table 1, Appendix VIII-1). About two-thirds of them in each setting were intellectually oriented. In contrast, authors' research orientation is related to their organizational affiliation. Sixty-seven per cent of authors in specialized agencies (or other non-academic organizations) were research oriented, followed by authors in universities (35 per cent) or colleges (35 per cent) and trailed by those in school systems (20 per cent) (Table VIII-2). Therefore, it is possible that orientation explains the relationship between organizational affiliation and quality. There are too few authors in specialized agencies to consider them separately while holding research orientation constant. Therefore, we had to combine them with authors in other organizations which probably lessened the relationship between being in a specialized agency and quality. Our data suggest that the relationship between working in a university and quality is partially explained by research orientation (Table VIII-3). Among the research oriented, 47 per cent of those in universities compared to 38 per cent of those in specialized agencies or other settings did good research. When research orientation is held constant, university affiliation is related primarily to better contribution to theory: 50 per cent of authors in universities contributed to theory compared to 29 per cent of those in specialized agencies or other organizations. When research oriented, authors in other organizations were just as likely to contribute to practice and use better methods of research as university-based authors.

TABLE VIII-2

ORIENTATION BY ORGANIZATIONAL AFFILIATION
(N=887)

Orientation	Organizational Affiliation at Time Research Paper was Written			
	Specialized Agency Other Organization	University	College	School System
Research	67%	35	35	20
Teaching or Administration	13	49	47	49
Service or Other	20	16	18	31
	100%	100%	100%	100%
	(61)	(700)	(71)	(55)

TABLE VIII-3

RESEARCH QUALITY BY ORGANIZATIONAL AFFILIATION BY ORIENTATION

Research Quality	Orientation								
	Research			Teaching or Administration			Service or Other		
	Organizational Affiliation at the Time of the Research Paper								
	University	College	Other	University	College	Other	University	College	Other
ABOVE AVERAGE contribution to theory	50% (90)	[39] (13)	29 (21)	29 (122)	11 (18)	13 (15)	19 (47)	- (4)	[0] (14)
ABOVE AVERAGE contribution to practice	41 (90)	[31] (13)	43 (21)	27 (122)	17 (18)	27 (15)	28 (47)	- (4)	[14] (14)
ABOVE AVERAGE use of research methods	49 (89)	[39] (13)	43 (21)	25 (122)	17 (18)	7 (15)	47 (47)	- (4)	[21] (14)

To summarize, authors in specialized agencies seem to have done the best research with respect to practice and methods, and university-based authors were almost as likely to have done good research on all dimensions. Authors in colleges seem not to have done as much good research as those in universities even when research oriented. This may be because they were isolated from research issues, and because that isolation impeded their research efforts.

School or Department[1]

As cited earlier, several observers have noted that service to school systems is emphasized to the detriment of research in schools of education. In view of this opinion, we wanted to examine the relationship between departmental affiliation and quality.

Since a number of behavioral science doctorates in our sample worked in schools of education, we controlled for any misleading effects this might have on the relationship between department and quality. We did this by relating departmental affiliation to quality, while holding constant type of doctorate. When type of doctorate is held constant, authors working in schools of education were as likely to do good research as those in liberal arts departments (Table VIII-4). The significant factor for research quality, then, seems to be the type of doctorate received by authors rather than the department in which they were working when they wrote their paper. Among behavioral science doctorates, 41 per cent of those in schools of education did good research compared to 40 per cent of those in behavioral science departments. Among education doctorates, 31 per cent of those in behavioral science departments compared to 29 per cent of those in education departments produced good research.

---

[1] Throughout our discussion we use the terms school or department interchangeably.

TABLE VIII-4

RESEARCH QUALITY BY DEPARTMENTAL AFFILIATION
BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate			
	Behavioral Sciences		Education	
	Departmental Affiliation			
	Behavioral Sciences	Education	Behavioral Sciences	Education
ABOVE AVERAGE contribution to theory	43 (54)	43 (21)	32 (22)	29 (147)
ABOVE AVERAGE contribution to practice	35 (54)	33 (21)	23 (22)	28 (147)
ABOVE AVERAGE use of research methods	41 (54)	48 (21)	38 (22)	29 (147)

There is one interesting exception to these tendencies, and that concerns the dimension of research methods. Authors working in departments different from the one in which they received their doctorate used better research methods than those working in the same type of department. Thus, 7 per cent more of those who received behavioral science doctorates and worked in a school of education used better methods than their doctoral colleagues in liberal arts departments.

Similarly, education doctorates in liberal arts departments were 9 per cent more likely to be using good research methods than education doctorates in schools of education. Perhaps there is something about working in a setting different from one's background that fosters clarification and improvement of research methods. We might speculate that constant questioning from people who may not share the same methodological assumptions and approaches sharpens one's use of research methods.

In Chapter V we noticed that most of the dissonant ratings of the three dimensions (i.e., those rated high on one dimension of quality and low on another) were made by judges who received their degrees in the behavioral sciences and then went to work in schools of education. That finding, together with the data here, supports the interpretation that a combination of different perspectives enhances one's use of research methods, and one's ability to be more sensitive to the requirements of good research.

Our aforementioned findings suggest that observers who lament the deleterious effect of schools of education on research quality should look one step earlier in the process. What they say about the service orientation of schools of education is probably true, but the data here suggest that this service orientation affects the research quality of people trained in schools of education but not the research of faculty members working there. Type of doctorate rather than departmental affiliation is what affects quality. As we have already noted in previous chapters, education doctorates are much more likely to have practice-oriented socialization, less likely to have graduate or career research experience, and thus less likely to be research oriented.

These data indicate that critics of educational research are correct, to a certain extent. The apparently practice-oriented nature of schools of education does not foster the development of researchers. Furthermore, by hiring primarily education graduates, schools of education seem to perpetuate a practice-oriented rather than a research-oriented climate, with adverse effects on quality. In brief, our data suggest that the generally lower quality of research produced in schools of education is explained by the type of doctorate of people hired to work in schools of education rather than by affiliation with a school of education *per se*.

Departmental Prestige

Because of selective hiring and presumably more stimulating intellectual interaction, we expect that people working in more prestigious university departments[1] will do better research than those working in less prestigious departments. This is true to a surprisingly small extent, however (Table VIII-5). Thirty-seven per cent of those in more prestigious departments did good research, compared to 31 per cent of those in less prestigious departments. One reason that this difference is not greater may be that those from less prestigious departments who publish are already somewhat atypical of those generally working there. Another explanation for the small relationship between departmental prestige and quality may stem from the fact that more education departments than discipline departments were evaluated as prestigious, which may distort the results reported in Table VIII-5. Therefore, we explore the relationship between departmental prestige and quality, while holding constant type of department (Table VIII-6).

---

[1] See Appendix VI-3 for a description of how departmental prestige was measured.

TABLE VIII-5

RESEARCH QUALITY BY PRESTIGE OF DEPARTMENT WHERE WROTE PAPER

Research Quality	Departmental Prestige	
	High	Other
ABOVE AVERAGE contribution to theory	39% (57)	32 (240)
ABOVE AVERAGE contribution to practice	35 (57)	30 (240)
ABOVE AVERAGE use of research methods	37 (57)	30 (240)

We see that departmental prestige is related to better research primarily in behavioral science rather than in education departments. In the behavioral sciences, departmental prestige is related to better theory and methods while in education prestige is related to contribution to practice. In the behavioral sciences, 57 per cent of authors in the more prestigious departments used good methods compared to 30 per cent of those in less prestigious departments, and 50 per cent

TABLE VIII-6

RESEARCH QUALITY BY DEPARTMENTAL PRESTIGE
BY DEPARTMENTAL AFFILIATION

Research Quality	Departmental Affiliation			
	Behavioral Sciences		Education	
	Departmental Prestige			
	High	Other	High	Other
ABOVE AVERAGE contribution to theory	[50%] (14)	34 (80)	35 (40)	32 (149)
ABOVE AVERAGE contribution to practice	[29] (14)	34 (80)	35 (40)	29 (149)
ABOVE AVERAGE use of research methods	[57] (14)	30 (79)	30 (40)	32 (149)

compared with 35 per cent contributed to theory. In the behavioral sciences, however, departmental prestige is negatively related to practice: 29 per cent of those in prestigious departments compared with 34 per cent in less prestigious departments contributed to practice. In schools of education, on the other hand, departmental prestige is related primarily to better practice: 35 per cent compared with 29 per cent in less prestigious departments contributed to practice. It is notable that those at prestigious schools of education did no better research than those in less prestigious behavioral science departments.

In sum, we have seen that departmental prestige is related to better theory and methods in the behavioral sciences, and to better contribution to practice in education departments. This association between prestige and the different dimensions of quality provides a clue to what is valued in the two fields. Education seems to value contribution to practice above all else. Because contribution to practice is related to prestige in education, we assume that is what is valued most; while in the behavioral sciences, contribution to practice seems to have a negative value, while contribution to theory and use of research methods are positively valued. By themselves these data are only directional, that is, they indicate the direction an interpretation might take. In the next chapter, however, we find additional evidence in support of this interpretation.

University Research Bureau Staff Membership

In Chapter VI we saw that working in a research bureau as a graduate student and being a research center staff member were positively related to better research. In this section we focus on university research staff members. We assume that research staff members have the chance to pool talents, skills, and ideas, as well as the opportunity for more research experience. Therefore we expect them to have done better research than independent faculty members. We anticipate this despite the findings of an earlier study by Persell.[1] In that study of authors in graduate schools of education, the quality of research done by research bureau staff members was compared with that of independent faculty members. Surprisingly, no difference was found in the quality of research done by the two groups of authors. There are two differences between the present study and the earlier one that lead us to expect different findings. First, this study includes behavioral scientists as well as education doctorates. Second, this study covers a much larger sample of authors (901 compared with 102). Differences between these two studies lead us to expect that our original hypothesis may be confirmed in this study, namely that authors in university research centers do better research than independent faculty members.

Our data indicate that research staff members produced better research than non-staff members, regardless of the type of research center in which they worked (Table VIII-7). Forty-three per cent of

---

[1] Persell, "Measuring Educational Research Quality."

TABLE VIII-7

RESEARCH QUALITY BY RESEARCH BUREAU STAFF STATUS BY ORGANIZATIONAL AFFILIATION

Research Quality	Organizational Affiliation					
	University		Specialized Agency or Other		College	
	Research Bureau Staff					
	Yes	No	Yes	No	Yes	No
ABOVE AVERAGE contribution to theory	43% (63)	33 (194)	25 (20)	10 (29)	- (3)	23 (31)
ABOVE AVERAGE contribution to practice	46 (63)	28 (194)	45 (20)	17 (29)	- (3)	23 (31)
ABOVE AVERAGE use of research methods	40 (62)	30 (194)	35 (20)	17 (29)	- (3)	26 (31)

university research bureau staff members and 35 per cent of specialized agency staff members are doing good research, compared to 30 per cent of independent university faculty. From now on, we will be concerned only with authors in a university research bureau, compared with independent faculty authors.

University research staff were slightly more likely than independent faculty to be intellectually oriented. Therefore it is possible that intellectual orientation explains why they do better research. Our data, however, indicate that intellectual orientation and research staff membership are independently related to better research. When analyzed jointly 55 per cent of intellectually oriented university research staff members did better research compared to 25 per cent of less intellectually oriented independent faculty (Table VIII-8).

Aside from being more intellectually oriented, university research staff were more research oriented than independent faculty as well. Consequently, we do not know if it is authors' research orientation or their research staff status that enhanced the quality of their research. It is also possible, of course, that research orientation and research staff status are independently related to quality, when considered together.

With respect to this question, our data reveal several striking findings. University research staff members did better work than independent faculty only when they were research oriented (Table VIII-9). Among university research staff, 57 per cent of the research oriented

TABLE VIII-8

RESEARCH QUALITY BY UNIVERSITY RESEARCH BUREAU STAFF STATUS BY INTELLECTUAL ORIENTATION

Research Quality	Intellectual Orientation					
	High		Medium		Low	
	University Research Bureau Staff?					
	Yes	No	Yes	No	Yes	No
ABOVE AVERAGE contribution to theory	59% (17)	48 (54)	36 (25)	28 (79)	40 (20)	26 (58)
ABOVE AVERAGE contribution to practice	53 (17)	30 (54)	48 (25)	28 (79)	40 (20)	26 (58)
ABOVE AVERAGE use of research methods	44 (16)	39 (54)	40 (25)	29 (79)	40 (20)	22 (58)

TABLE VIII-9

RESEARCH QUALITY BY UNIVERSITY RESEARCH CENTER STAFF STATUS BY ORIENTATION

Research Quality	Orientation					
	Research		Teaching or Administration		Service or Other	
	University Research Staff Member					
	Yes	No	Yes	No	Yes	No
ABOVE AVERAGE contribution to theory	58% (33)	47 (55)	16 (19)	32 (100)	- (11)	11 (36)
ABOVE AVERAGE contribution to practice	58 (33)	33 (55)	26 (19)	27 (100)	- (11)	22 (36)
ABOVE AVERAGE use of research methods	56 (32)	46 (55)	21 (19)	26 (100)	- (11)	19 (36)

compared to 21 per cent of authors who were not research oriented did better research. By itself being a staff member of a university research center did not significantly enhance research quality. In order for this structural arrangement to facilitate better research, authors needed to be research oriented (i.e., prefer research first as a professional activity). In short, the strong positive relationship between university research staff status and better research is generally contingent upon being research oriented.

Research orientation is related to better research even when research center staff membership is held constant. That relationship is sharply increased, however, among research center staff members. Among the research oriented, 57 per cent of staff members did better research, compared to 42 per cent of non-staff. This suggests that research oriented authors who worked in a facilitative setting did better research than those who were research oriented but did not work in such a setting.

We found earlier that work in a research center as a graduate student was related to better research only when it occurred in a prestigious department. The question now arises of whether the relationship between being a university research staff member and quality is contingent upon high departmental prestige. Our data indicate that it is not. Research staff members produced better research, regardless of the prestige of the department in which they worked (Table VIII-10). Moreover, research staff membership seems largely to explain the relationship between departmental prestige and quality. That is, the

TABLE VIII-10

RESEARCH QUALITY BY UNIVERSITY RESEARCH CENTER STAFF STATUS BY DEPARTMENTAL PRESTIGE

Research Quality	Departmental Prestige Where Paper Was Written			
	High		Other	
	University Research Staff Member			
	Yes	No	Yes	No
ABOVE AVERAGE contribution to theory	45% (20)	35 (37)	42 (43)	32 (156)
ABOVE AVERAGE contribution to practice	40 (20)	32 (37)	48 (43)	27 (156)
ABOVE AVERAGE use of research methods	45 (20)	32 (37)	38 (42)	29 (156)

tendency of authors in more prestigious departments to have been research center staff members seems to explain why departmental prestige is related to quality. When authors in less prestigious departments were research center staff members, they were just as likely as those in more prestigious departments to do better research. Among research staff, 42 per cent of those in more prestigious departments did better research, compared to 43 per cent of those in less prestigious departments. (Among independent faculty, 31 per cent of those in more prestigious departments compared with 28 per cent of those in less prestigious departments produced better research.)

In sum, being a staff member of a university research bureau enhanced the quality of work done by research oriented authors. When authors were research oriented and also staff members of research centers, nearly 60 per cent did good work. Clearly research centers serve a very important facilitative function. Research staff did better research than independent faculty even when intellectual orientation and research orientation were held constant. Staff members excelled at research, however, only when research oriented. If they lacked that orientation, they were less likely to do good research than independent faculty. In short, university research centers may facilitate better research through two processes. First, work in a research center as a graduate student increases career research experience, as noted in Chapter VI. Second, being a university research center staff member enhanced the quality of work done by research oriented authors.[1]

---

[1] For a more extensive discussion of why research centers facilitate research, see Sieber, Reforming the University.

Summary

In this chapter we have seen that several research settings seem to affect quality, independently of research orientation and intellectual orientation. University-based authors generally did better research than those in colleges, school systems, or other settings, although this is explained on all dimensions except theory by their greater research orientation. While the number of examples is too small to be conclusive, authors in specialized agencies also seem to be doing better research than those in colleges, school systems, or other organizations, albeit this may be due to their substantially greater tendency to be research oriented. These findings suggest that the issue of whether research should be done in universities or in specialized agencies is somewhat spurious. The real differences in research quality seem to be related to working in a university or specialized agency, on the one hand, compared to working in a college, school system, or another organization, on the other.

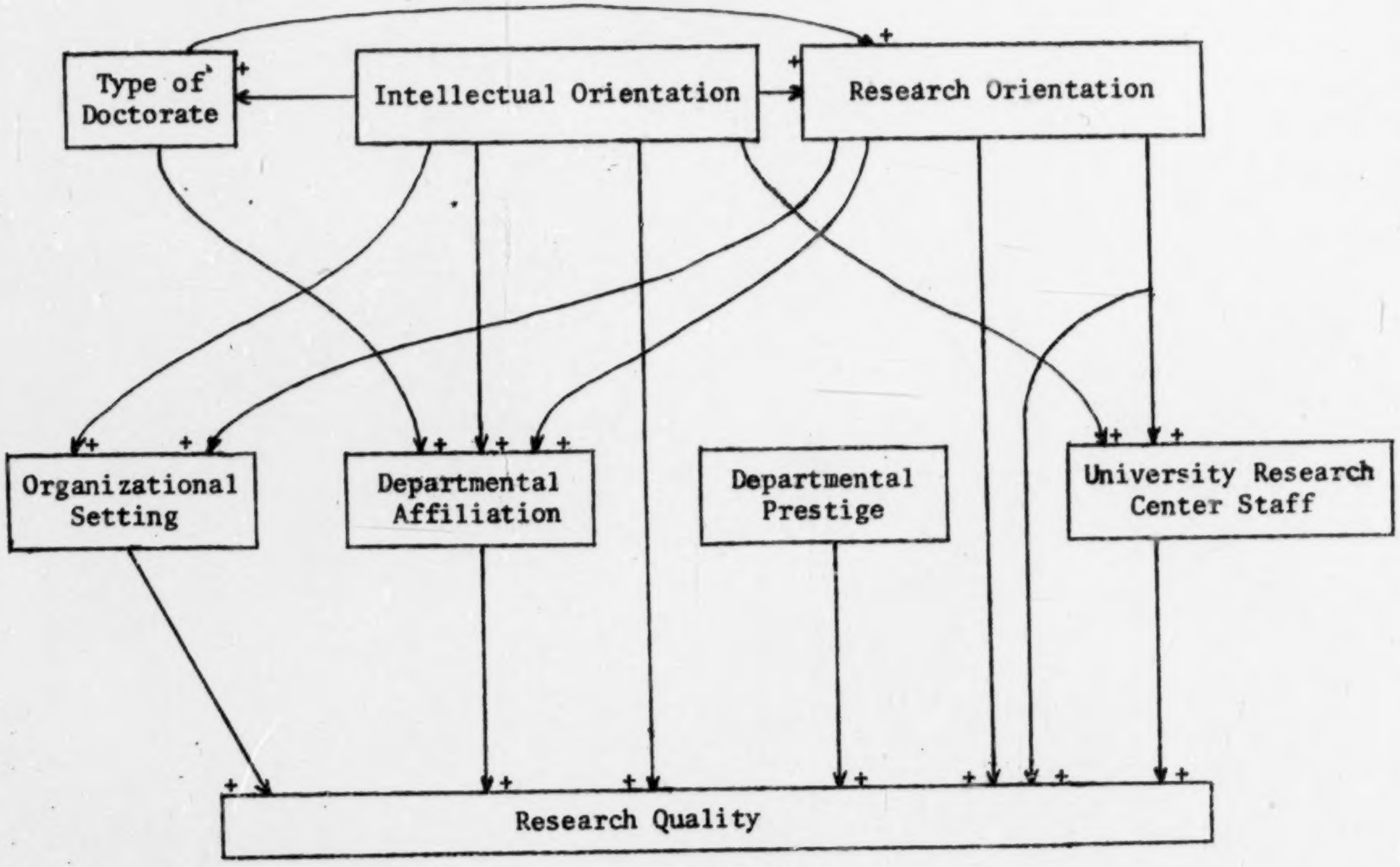
In addition, we considered the issue of whether working in a school of education was detrimental for research. We discovered that type of doctorate rather than departmental affiliation seems to explain why authors in behavioral science departments did better research than did authors in schools of education. These data suggest that graduate training may be more important for quality than the departmental setting in which the work was done.

We also explored the assumption that authors in more prestigious departments do better research than those in less prestigious departments, and found that they did, although the difference is surprisingly small. Prestige is more strongly related to better research (specifically contribution to theory and use of research methods) in the behavioral sciences than in education. In education, departmental prestige seems to be more highly related to contribution to practice than to theory or methods.

Finally, we examined the association between a university research center and quality, and found that center staff did better research than independent faculty, even when intellectual orientation was held constant. When research orientation was held constant, however, we found that research bureau staff did better research than independent faculty only when they were research oriented. We also considered whether the relationship between research staff status and better research was contingent upon departmental prestige, and found that it was not. Instead, research staff status seemed to explain the relationship between departmental prestige and quality. Apparently the tendency of authors in prestigious departments to have been staff members of research bureaus explains why they did better research than authors in less prestigious departments. When authors in less prestigious departments were staff members of research centers, they were just as likely to do good research as those in more prestigious departments.

FIGURE VIII-1

DOCUMENTED INTERRELATIONSHIPS BETWEEN INTELLECTUAL ORIENTATION, TYPE OF DOCTORATE, RESEARCH ORIENTATION, RESEARCH SETTINGS, AND QUALITY


Type of Doctorate
Intellectual Orientation
Research Orientation
Organizational Setting
Departmental Affiliation
Departmental Prestige
University Research Center Staff
Research Quality
+

(no line between two variables means no relationship)

CHAPTER IX

THE REWARD AND COMMUNICATION SYSTEMS

In the preceding analysis, we considered the background, role socialization, and work settings of authors of recent educational research in an effort to understand how these variables are related to better research. In this chapter we widen our perspective and consider the larger social system in which the authors work. The way the reward system operates in the behavioral sciences and in education may be related to the quality of work produced in the two fields. The importance of the reward system in science has been stressed by such sociologists of science as Cole, Merton, Storer, and Zuckerman.[1] We will refer to their work where relevant throughout this chapter.

The Operation of the Reward System

Rewards in a field may be allocated by several processes. Zuckerman notes the following mechanisms for distributing rewards in science: selective recruitment and socialization, allocation of jobs, allocation of resources, differential access to publication, and allocation of honorofic awards.[2] In this study we measure three of these five processes: access to publication, allocation of resources (research grants), and the allocation of positions.

---

[1] Cole, "Social Structure of Science"; Robert K. Merton, "The Matthew Effect in Science," Science, January 5, 1968, pp. 56-63; Norman W. Storer, The Social System of Science (New York: Holt, Rinehart, and Winston, 1966); Harriet A. Zuckerman, "Stratification in American Science," Sociological Inquiry, XL (Spring), 235-57.

[2] Zuckerman, "Stratification in American Science."

Access to Publication

One reward the social system of research offers is having one's work published. If access to publication is selective and if acceptance for publication is based on a paper's quality, then we would expect to find a positive relationship between quality and published output. In physics more productive authors are likely to do better research than less productive ones.[1] Several studies have assumed that this relationship exists in educational research as well, and they have therefore analysed antecedent conditions (e.g., training or communication patterns) in relation to productivity rather than in relation to quality. To verify this assumption, and to understand the operation of one aspect of the reward system in educational research, we devised a measure of productivity based on the average number of research articles published per year since the author's first publication.[2] (See Appendix IX-1 for a further description of how this index of productivity was formed.) Surprisingly, productivity is positively related to only one dimension of research quality, namely contribution to theory (Table IX-1). Thirty-eight per cent of highly productive authors are contributing to theory, compared to 25 per cent of the less productive. Productivity is virtually unrelated to contribution to practice or use of research methods. Moreover, the association that exists between productivity and contribution to theory occurs only among authors in the behavioral sciences (Table IX-2). Among behavioral science doctorates, 54 per cent of the high producers are contributing to theory compared to 36 per cent of low producers. Otherwise there is no relationship between quality and productivity. These data suggest that, with the exception of contribution to theory in the behavioral sciences, frequent publication does not seem to be related to better research. Later in this chapter we will return to the question of the likely consequences of a system where being rewarded (in this case being published frequently) is unrelated to quality.

---

[1] Cole, "Social Structure of Science."

[2] Blaine Worthen, of the Laboratory of Educational Research, University of Colorado, has verified the method of asking respondents to indicate how many research papers they have published. In his study of research assistanceships, he asked respondents to report the number of publications they had produced. Then he asked for a bibliography of their publications. A sample of these bibliographies was checked by consulting the original journals. He found that very few authors reported more articles than actually had been published. This information was reported to us by Blaine Worthen, in a conversation.

The non-association between productivity and quality in education has methodological significance for studies that have examined the correlates of research productivity. Examination of factors which affect productivity is interesting and important in its own right but it is not identical with explanations of quality. The existence of relationships between background, training, or career experiences and productivity may or may not indicate relationships between those characteristics and quality.

Allocation of Research Grants

Perhaps the absence of a relationship between productivity and

TABLE IX-1

RESEARCH QUALITY BY PRODUCTIVITY (AVERAGE ANNUAL NUMBER OF ARTICLES)

Research Quality	Average Annual Number of Research Articles		
	< 1	< 2	2 or more
ABOVE AVERAGE contribution to theory	25% (98)	33 (82)	38 (94)
ABOVE AVERAGE contribution to practice	29 (98)	31 (82)	32 (94)
ABOVE AVERAGE use of research methods	28 (98)	35 (82)	33 (94)

TABLE IX-2

RESEARCH QUALITY BY PRODUCTIVITY BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate					
	Behavioral Sciences			Education		
	Average Annual Number of Articles					
	< 1	< 2	2 or more	< 1	< 2	2 or more
ABOVE AVERAGE contribution to theory	36% (28)	32 (31)	54 (28)	26 (66)	38 (49)	24 (50)
ABOVE AVERAGE contribution to practice	39 (28)	39 (31)	39 (28)	29 (66)	31 (49)	24 (50)
ABOVE AVERAGE use of research methods	43 (28)	29 (31)	43 (28)	32 (66)	27 (49)	34 (50)

quality in education is simply one exception in an otherwise efficient reward system. If so, we would expect that other rewards, such as receiving a research grant,[1] would be related to better research. We asked authors,

> Have you ever applied for a research grant from the U.S.O.E., N.I.M.H., Ford, Rockefeller, or Carnegie foundations, or any other such organization that supports research?
>
> IF YES, have you ever received a research grant from such an organization?

Fifty-nine per cent of respondents applied for research grants, and 70 per cent (398) of those received a grant.

Among those who applied for grants, it is surprising that better researchers are no more likely to have received research grants than poorer researchers, in either the behavioral sciences or in education (Table IX-3), with one exception. In the behavioral sciences, authors whose papers were rated better with respect to research methods are somewhat more likely to have received a research grant during their careers: 82 per cent of those rated above average with respect to research methods received grants, compared to 71 per cent of those rated below average. In education, 75 per cent of average researchers have received a research grant, compared to 66 per cent of better researchers.

---

[1] We chose this measure of being funded rather than the one taken from the question, "Did the project (on which the paper is based) receive financial support for research expenses from university funds, a foundation, a governmental agency, etc.?" because here we are interested in being funded as an indicator of the operation of the reward system. In other connections it is useful to explore the quality of particular papers in terms of whether or not they were based on research that was funded.

TABLE IX-3

APPLIED FOR AND RECEIVED A RESEARCH GRANT BY QUALITY BY TYPE OF DOCTORATE

	Type of Doctorate					
	Behavioral Sciences			Education		
Rated Quality of Contribution to Theory						
	Above Average	Average	Below Average	Above Average	Average	Below Average
Applied for and received a research grant	81% (36)	70 (20)	81 (21)	60 (37)	83 (36)	60 (57)
Rated Quality of Contribution to Practice						
Applied for and received a research grant	81 (32)	74 (34)	82 (11)	68 (40)	69 (45)	62 (45)
Rated Quality of Use of Research Methods						
Applied for and received a research grant	82 (33)	78 (23)	71 (21)	69 (42)	74 (43)	56 (45)

Allocation of Jobs

Another measure of how the reward system operates in a field may be seen in the relationship between student or faculty status and quality, and that between academic rank and quality, insofar as we would expect that faculty status and high academic rank would be related to better research. However, it must be borne in mind that these relationships might be contaminated by generational differences. For example, we have already noted in Chapter VII that younger authors do better research. We will briefly consider the relationship between position and quality and that between rank and quality. Both student or faculty status and academic rank were measured at the time authors wrote the paper that was rated.

Student or Faculty Status

One of the assumptions of the research enterprise is that gaining experience and stature is related to better work. In view of this, we expected that the research of university faculty members would be considered better than that of students. We measured student status by asking: "Was the paper part of your master's thesis, doctoral dissertation, or a student paper?" For 37 per cent of respondents the paper was done as part of their student work.[1] University faculty status was determined by asking authors, "At the time of this research, what was your main institutional affiliation?" Among those who indicated that they were affiliated with a university, all those who had not written the paper as part of their student work were considered faculty in this table.

---

[1] We chose to measure student status with this question rather than with the one asking for the respondent's "principal position" at the time he did his research, since many fewer indicated that they were students in response to that question.

In education, university faculty did better research than students as expected: 35 per cent of faculty did better research compared to 27 per cent of students. In the behavioral sciences, however, 47 per cent of students compared to 38 per cent of faculty did better research (Table IX-4). Perhaps older behavioral scientists who did research on education are not as good researchers as the students now venturing into educational research from the behavioral sciences.

Academic Rank

With respect to academic rank as well, we assumed that authors with higher academic ranks would have done better research than those of lower rank. We determined academic rank by asking authors to indicate, "What was your principal position at this institution?" (i.e., the one they were affiliated with at the time they wrote their paper). For those in colleges or universities, we coded the academic rank they indicated.[1]

Surprisingly, we found that authors of lower academic rank did better research than full professors, in both education and the behavioral sciences, except with respect to contribution to educational practice. In the behavioral sciences, 50 per cent of assistant

---

[1] The code categories were: professor, associate professor, assistant professor, lecturer or instructor, research associate or researcher, research assistant, student, other.

TABLE IX-4

RESEARCH QUALITY BY STUDENT OR FACULTY
STATUS BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate			
	Behavioral Sciences		Education	
	Status			
	University Faculty	Student	University Faculty	Student
ABOVE AVERAGE contribution to theory	45% (51)	43 (21)	35 (78)	28 (76)
ABOVE AVERAGE contribution to practice	33 (51)	48 (21)	33 (78)	26 (76)
ABOVE AVERAGE use of research methods	35 (51)	48 (21)	36 (78)	28 (75)

professors[1] were rated well on theory and research methods compared to 36 per cent of full professors. In education, 38 per cent of assistant professors compared to 24 per cent of full professors contributed to theory and used good methods (Table IX-5).

There are at least two interpretations of these findings. Either better researchers are not being promoted in rank, or there are a number of younger authors who are doing better work. Our evidence suggests that it is the younger authors among those of lower academic rank who are doing better research (see Table 1, Appendix IX-2). Therefore the question arises, has the system rewarded these good young researchers, and thereby increased their desire to do research?[2] Specifically, are the better researchers more likely to have been promoted in rank or to have moved from a less prestigious to a more prestigious department between the time they wrote their papers (which were rated by the judges in this study) and the time they answered the questionnaire? We will consider each of these possibilities.

Career Mobility

Promotion in Academic Rank

We can analyse the operation of the reward system more clearly by measuring whether authors have been promoted in academic rank since they wrote their papers. Are the authors whose papers were rated

---

[1] Although the number of assistant professors is very small, when it is combined with associate professors the result is similar.

[2] In a later section of this chapter we consider the evidence bearing on the relationship between being rewarded and preferring research as a professional activity.

TABLE IX-5

RESEARCH QUALITY BY ACADEMIC RANK BY TYPE OF DOCTORATE

Research Quality	Type of Doctorate									
	Behavioral Sciences					Education				
	Academic Rank									
	Prof.	Assoc. Prof.	Ass'nt Prof.	Lect'r/ Inst'r	Res. Assoc.	Prof.	Assoc. Prof.	Ass'nt Prof.	Lect'r/ Inst'r	Res. Assoc.
ABOVE AVERAGE contribution to theory	38% (24)	54 (13)	54 (13)	- (1)	[36] (14)	25 (32)	25 (24)	36 (33)	27 (15)	[50] (10)
ABOVE AVERAGE contribution to practice	42 (24)	31 (13)	23 (13)	- (1)	[50] (14)	25 (32)	21 (24)	27 (33)	13 (15)	[50] (10)
ABOVE AVERAGE use of research methods	33 (24)	54 (13)	46 (13)	- (1)	[43] (14)	22 (32)	21 (24)	39 (33)	14 (14)	[60] (10)

better more likely to have been promoted? As noted earlier, we measured academic rank by asking authors, "What was your principal position at this institution? (e.g., professor of educational psychology)" at the time of this research. Promotion in academic rank was determined by comparing academic rank at the time the paper was written with academic rank at the time the questionnaire was returned.[1]

In education, mediocrity seems to be the key to promotion in academic rank (Table IX-6). Of those who did average research, 61 per cent have since been promoted in rank, compared to only 44 per cent of those who did better research. These findings suggest that the promotion system in education is not rewarding researchers whose work was rated better. If sociologists of science are correct in believing that the operation of the reward system serves to establish models of outstanding work, as well as to reinforce the activities of those whom it rewards, then the system is operating in education in such a way as to discourage better research.

In the behavioral sciences there are too few cases to be conclusive, but the data suggest that having used better research methods is positively related to promotion in rank. Among behavioral scientises, 35 per cent of authors rated well with respect to research methods were promoted in rank, compared to 17 per cent of researchers whose papers were considered below average on that dimension (Table

---

[1] We had a limited sample for whom we could make these comparisons, since promotion in academic rank applies only to college or university personnel. We excluded researchers since they were not considered to have positions that were comparable to other academic positions.

TABLE IX-6

CHANGE IN ACADEMIC RANK BY QUALITY
BY TYPE OF DOCTORATE

Change in Academic Rank	Type of Doctorate					
	Behavioral Sciences			Education		
	Rated Quality of Contribution to Theory					
	Above Average	Average	Below Average	Above Average	Average	Below Average
Stayed up--full professor	[35]	[36]	[55]	20	21	27
Moved up	[24]	[14]	[27]	36	68	44
Stayed the same	[35]	[43]	[9]	32	4	27
Can't tell	[6]	[7]	[9]	12	7	3
	100%	100%	100%	100%	100%	101%
	(17)	(14)	(11)	(25)	(28)	(34)
	Rated Quality of Contribution to Practice					
Stayed up--full professor	[58]	32	-	20	21	27
Moved up	[17]	18	-	52	59	39
Stayed the same	[8]	50	-	12	17	30
Can't tell	[17]	0	-	16	3	3
	100%	100%		100%	100%	99%
	(12)	(22)	(8)	(25)	(29)	(33)
	Rated Quality of Use of Research Methods					
Stayed up--full professor	[29]	[39]	[58]	18	25	27
Moved up	[35]	[8]	[17]	46	57	43
Stayed the same	[29]	[46]	[17]	25	11	27
Can't tell	[6]	[8]	[8]	11	7	3
	99%	101%	100%	100%	100%	100%
	(17)	(13)	(12)	(28)	(28)	(30)

IX-6). With respect to theory, however, authors of below average papers were more frequently promoted than authors of better papers, although there are too few examples to be conclusive. For behavioral scientists, contribution to educational practice seems unrelated to promotion, although there are too few cases to be sure.

From the above evidence it appears that authors of better educational research papers are no more likely to be promoted than authors of mediocre or poor papers. Before we adopt this conclusion, however, let us consider another indication of how the reward system operates. We turn now to interdepartmental movement, particularly changes in departmental prestige[1] between the time the research was done and the time the questionnaire was answered.

Among education doctorates, authors of better research papers are only slightly more likely than authors of average or poor papers to have moved from a less prestigious department to a more prestigious department (when the three dimensions are averaged, 9 per cent of authors of better papers compared to 3 per cent of authors of poorer papers made such a move). (Table IX-7.)

So far, we have considered only authors going from less prestigious to more prestigious departments. Continuing to work in a prestigious department may also be conceived of as a reward the system has to offer. When we consider those who were at a prestigious department when they wrote their paper and who either stayed there or moved to

---

[1] The prestige of education departments was based on Sieber's survey and the prestige of behavioral science departments was based on Cartter's evaluations. See Appendix IX-3 for a more detailed discussion of the index of mobility in departmental prestige.

TABLE IX-7

MOBILITY IN DEPARTMENTAL PRESTIGE BY QUALITY BY TYPE OF DOCTORATE

Change in Departmental Prestige	Type of Doctorate					
	Behavioral Sciences			Education		
Quality of Contribution to Theory						
	Above Average	Average	Below Average	Above Average	Average	Below Average
Moved from less to more prestigious	14%	13	0	9	7	2
Stayed same - prestigious	4	6	0	13	12	5
Stayed same - less prestigious	72	75	85	30	42	37
Moved from more to less prestigious	10	6	15	49	39	57
	100% (29)	100% (16)	100% (20)	101% (47)	100% (41)	101% (65)
Quality of Contribution to Practice						
Moved from less to more prestigious	8	11	[23]	9	2	5
Stayed same - prestigious	4	0	[8]	14	10	5
Stayed same - less prestigious	76	78	[62]	33	43	32
Moved from more to less prestigious	12	11	[8]	44	45	58
	100% (25)	100% (27)	(13)	100% (43)	100% (51)	100% (59)
Quality of Use of Research Methods						
Moved from less to more prestigious	16	5	5	9	6	2
Stayed same - prestigious	8	0	0	16	10	3
Stayed same - less prestigious	72	80	80	24	42	39
Moved from more to less prestigious	4	15	15	51	42	56
	100% (25)	100% (20)	100% (20)	100% (45)	100% (48)	100% (59)

another equally prestigious department, we find a positive relationship between quality and continuing at a prestigious department. In education, 14 per cent of better researchers stayed at prestigious departments compared to 4 per cent of poorer researchers. Finally, 57 per cent of poorer researchers compared to 48 per cent of better ones have moved from a more prestigious department to a less prestigious one. Thus, having done poor research is related to downward mobility in terms of departmental prestige.

We may speculate that upward mobility requiring a change in department may be explained more by factors other than the rated quality of one's published research (such as productivity, reputation, grantsmanship, teaching ability, or something else), whereas quality may be related to remaining at a prestigious department because an author's work may be more visible within the same or a similar department.

In the behavioral science, by way of contrast, better research (with respect to theory and methods) is somewhat more strongly related to moving from a less prestigious to a more prestigious department (Table IX-7). None of the authors whose papers were rated below average with respect to theory moved from a less prestigious to a more prestigious department, while 14 per cent of those who contributed to theory did so. Similarly, only 5 per cent of authors whose papers were rated below average with respect to their use of research methods moved up, compared to 16 per cent of those rated above average. Although there are too few cases to be conclusive, our data suggest that contribution to educational practice is not related to upward mobility in departmental prestige in the behavioral sciences.

When we consider behavioral science authors who continued at a prestigious department, we see that being rated above average with respect to research methods is related to staying in a more prestigious department; 8 per cent of the highly rated continued, compared to none of those rated average or below. Contribution to theory has an inconclusive relationship to remaining at a more prestigious department.

To summarize, we have seen that in education better researchers are either no more likely to be rewarded than poorer researchers, or they are only slightly more likely to be rewarded, depending on the measure used. We measured rewards in terms of being published, receiving a research grant, being promoted in academic rank, or moving from a less prestigious to a more prestigious department. In the behavioral sciences, on the other hand, we have seen that authors of papers rated better on one or more dimension of quality are somewhat more likely than authors of lower rated papers to be rewarded: contribution to theory is related to productivity, using better methods is related to receiving a research grant, and better theory and methods are related to promotion in academic rank and to moving from a less prestigious to a more prestigious department. We conclude that in the behavioral sciences the reward system seems somewhat more likely to reinforce better research than it does in education.

Consequences of the Way the Reward System Operates

Given this evidence of how the reward system operates in the two fields, what are the likely consequences for the research enterprise? As we mentioned earlier, sociologists of science[1] have noted that the reward system serves two key functions in science. First, it provides examples of what is valued in a social system. Second, being rewarded is presumed to reinforce the motivation of scientists, thereby encouraging them in their efforts. Glaser found that the more scientists were rewarded, the more motivated they were to do research, and the more productive they were.[2] Therefore, we wondered whether authors who had been rewarded[3] were more likely to prefer research (i.e., be research oriented) than authors who had not been rewarded. Because preference for research and quality are strongly related, we controlled for quality while relating an index of having been rewarded to preference for research (i.e., research orientation).

As we already know, quality is related to research orientation (preference). What is of interest here is the strong independent relationship between being rewarded and preferring research (Table IX-8). Furthermore, better researchers are even more likely than poorer

---

[1] See, for example, Cole, "Social Structure of Science"; Merton, "The Matthew Effect in Science"; Storer, The Social System of Science; Zuckerman, "Stratification in American Science."

[2] Barney G. Glaser, Organizational Scientists: Their Professional Careers (Indianapolis: Bobbs Merrill, 1964), pp. 18-25.

[3] Authors were classified as having been rewarded if they had one or more of the following: if they had published on the average more than two papers per year since their first publication, if they had received a research grant, if they had been promoted in academic rank between the time they wrote the paper and the time they answered the questionnaire, or if they had moved from a less prestigious to a more prestigious department during that time period. See Appendix IX-4 for a more detailed description of how this index was constructed.

TABLE IX-8

RESEARCH ORIENTATION BY BEING REWARDED BY RESEARCH QUALITY

(per cent research oriented)

Research Quality	Rewarded	
	Yes	No
Contribution to theory:		
Above average	62% (68)	38 (37)
Average	42 (64)	16 (32)
Below average	28 (87)	21 (57)
Contribution to practice:		
Above average	52 (69)	39 (36)
Average	45 (80)	21 (44)
Below average	30 (70)	17 (46)
Use of research methods:		
Above average	61 (72)	41 (34)
Average	40 (72)	22 (41)
Below average	27 (75)	14 (50)

researchers to prefer research if they have been rewarded. Among better researchers, 58 per cent (when we average the three dimensions) of the rewarded ones prefer research compared to 36 per cent of the unrewarded ones. These data provide evidence that the reward system does serve to enhance authors' preference for research. At every level of quality, authors who were rewarded are more likely to prefer research than authors who were not rewarded. In view of these findings, it is possible that authors of well rated research papers may tend to lose interest in research if they are not recognized and rewarded for their accomplishments. Furthermore, if they see mediocre researchers being rewarded as frequently as better researchers, then it is not hard to understand how they might lose interest in doing research and direct their efforts toward other activities.

Interpretations of Why the Reward System Operates the Way It Does

"Publish or Perish"

Since the reward system seems to have serious consequences for the research enterprise, we want to consider its operation somewhat further. If, as we have seen, quality is relatively unrelated to rewards in education, it may be that rewards are allocated according to the *quantity* of an author's output. If that were the case, the "publish or perish" axiom would be true. The evidence in support of this interpretation, however, is mixed.

Productivity is related to having been funded sometime in one's career (Table IX-9), especially in the behavioral sciences. In the

TABLE IX-9

APPLIED FOR AND RECEIVED A RESEARCH GRANT BY PRODUCTIVITY BY TYPE OF DOCTORATE

	Type of Doctorate					
	Behavioral Sciences			Education		
	Productivity -- Average Annual Number of Research Articles Since First One					
	< 1	< 2	2 or More	< 1	< 2	2 or More
Percent who Applied for and received research grant	65 (43)	88 (69)	82 (68)	66 (105)	63 (81)	75 (93)

behavioral sciences, the very productive are 17 per cent more likely than the less productive to have received a research grant, while in education they are 9 per cent more likely. We cannot tell if receiving a research grant encouraged publication (e.g., by freeing time for research and writing and by requiring production of one or more reports) or if the very productive are more likely to receive research grants. Probably the two are mutually reinforcing.

In education, more productive authors are 6 per cent more likely to have moved from a less prestigious to a more prestigious department, whereas in the behavioral sciences they are only 3 per cent more likely to have moved from a less prestigious to a more prestigious department (Table IX-10). However, productivity is negatively related to promotion in rank in both education and the behavioral sciences (Table IX-11). The highly productive are 6 per cent less likely to have been promoted in rank in both fields. In short, in education productivity is related to receiving a research grant and slightly to upward mobility in departmental prestige. While this is not overwhelming evidence that productivity is the key to being rewarded in education, it does suggest that productivity is more related to rewards than is quality.

In the preceding analysis we have seen that quality is unrelated to being rewarded in education, and that productivity is somewhat related to being rewarded in education, but not so strongly that the "publish or perish" interpretation of how rewards are allocated is convincingly upheld. There are at least two other interpretations of

TABLE IX-10

MOBILITY IN DEPARTMENTAL PRESTIGE BY
PRODUCTIVITY BY TYPE OF DOCTORATE

Index of Mobility in Departmental Prestige	Type of Doctorate					
	Behavioral Sciences			Education		
	Productivity -- Average Annual Number of Research Papers Since First One					
	< 1	< 2	2 or More	< 1	< 2	2 or More
Moved from less to more prestigious	9%	7	12	3	8	9
Stayed same -- prestigious	0	0	7	6	6	8
Stayed same -- less prestigious	81	75	71	42	36	38
Moved from more to less prestigious	11	19	10	50	50	45
	101%	101%	100%	101%	100%	100%
	(57)	(59)	(59)	(127)	(99)	(100)

TABLE IX-11

CHANGE IN ACADEMIC RANK BY PRODUCTIVITY
BY TYPE OF DOCTORATE

Index of Change in Academic Rank	Type of Doctorate					
	Behavioral Sciences			Education		
	Productivity -- Average Annual Number of Research Papers Since First One					
	< 1	< 2	2 or More	< 1	< 2	2 or More
Stayed the same - full professor	31%	22	8	23	23	28
Moved up	31	32	25	49	45	43
Stayed the same	31	32	25	19	23	25
Can't tell	6	11	8	9	8	4
	99%	100%	101%	100%	99%	101%
	(32)	(37)	(40)	(79)	(60)	(53)

why better research is not rewarded in education to the extent that it is in the behavioral sciences and physics.[1] First, as mentioned in Chapter III, the communication system may operate so inefficiently in education that it greatly impedes the effective operation of the reward system. Second, the field of education may consider other activities as important as research, with the result that contributions in other endeavors are rewarded independently of research quality. We will briefly consider each of these possibilities.

The Nature of the Communication System

There is considerable evidence to indicate that the communication system in educational research is very diffuse. Nelson[2] compared the diffusion rates of papers presented at the 1968[3] American Educational Research Association (AERA) with those of nine professional and scientific associations in other fields. He found that AERA paper-givers submit their papers to a far wider array of journals than do people in any other professional association. Nelson found that half of the responding authors of contributed papers had submitted their papers for publication. But these 102 authors submitted their manuscripts to 64 different journals, "only four of which published or accepted more than two of these manuscripts each!"[1] As Nelson notes, "A person would have to read a minimum of 12 different journals to read one half of the [contributed papers] which were published or accepted for publication within one year after the AERA meeting."[2] Furthermore, he reports,

---

[1] The relationship in physics was documented by Cole, "Social Structure of Science."

[2] Carnot E. Nelson (with the assistance of William Garvey and Nan Lin), "The Postmeeting Journal Dissemination of Material Presented at the 1968 American Educational Research Association Annual Meeting" (paper presented at the 1970 American Educational Research Association Annual Meeting in Minneapolis, Minnesota, March, 1970). (Mimeographed.)

[3] Nelson, Garvey and Lin studied the same AERA paper-givers as the ones inclueed in our sample of recent educational researchers.

> . . . 47 AERA symposium authors submitted manuscripts to 34 different journals. In sociology, which has the next highest rate of diffusion, the large number of journals in which papers appear is explained by the fact that the first submissions by many sociologists were to a few very prestigious journals which rejected most of the manuscripts. These rejected manuscripts were then resubmitted to a wide variety of journals.[3]

Among AERA paper-givers, however, the "diffusion in journal outlets began with the first submittion. . . . Therefore, it appears that there is no one journal or small group of journals to which most educational researchers will submit their work."[4] This evidence for the absence of core journals in education suggests that even within the single largest professional association in educational research, authors do not have common means of communicating with each other beyond attending the annual AERA meeting.

When authors in a field do not share their research by submitting it to a common core of journals, several consequences are likely:

(a) authors are less likely to develop and share critical standards of quality;

(b) better research is less likely to be visible throughout the field (see Chapter III for a discussion of the nature and importance of visibility);

---

[1] Nelson, "The Postmeeting Journal Dissemination," p. 4.

[3] Ibid. [2] Ibid., p. 7. [3] Ibid., p. 8.

(c) researchers are less likely to receive "competent response"[1] from relevant colleagues in the field;

(d) the reward system is less likely to operate effectively (i.e., to reward those doing the best research wherever it appears) because so much research would be unknown.

Not only are the publication patterns of AERA paper-givers more diffuse than those of nine other scientific or professional association members, but 53 per cent of the educational researchers in this sample do not belong to AERA. Furthermore, the 901 authors in this sample belong to 417 different professional associations. Moreover, they published their papers in 113 different journals. Not surprisingly, judges in this study recognized[2] only 25 per cent of even the best research papers and only 7 per cent of all papers. Thus, even the best research suffers from low visibility in this highly diffuse communication system. This means that the inefficient communication processes among AERA authors reported by Nelson probably understate the problem for educational research at large.

In short, judged by the scattered publication patterns of AERA paper-givers, the spread of membership in professional associations among this sample, the array of journals in which papers in this sample were published, and the low proportion of papers that were recognized by judges, we conclude that the communication system in educational research is highly diffuse, and therefore cannot insure visibility of the best research throughout the system.

---

[1] The importance of "competent response" is discussed in detail by Norman Storer in The Social System of Science.

[2] As noted in Chapter IV, we measured the judge's recognition of papers with three questions on the Specification Form. In this case, recognition is based on the question, "Do you remember having seen this paper before?" Of the three this was the question that the most judges answered affirmatively.

Cole[1] has discussed the interdependence of the reward and communication systems in modern physics. If we assume the interdependence of the two in educational research, and if we assume that the communication system does not insure the visibility of good research as we have just suggested, then we have one plausible explanation of why the reward system operates the way it does in education.

The Existence of Multiple Goals

Another interpretation of why the reward system operates the way it does may reside in the existence of multiple goals in education. In our analysis throughout the second half of this volume we have noted that the context in which educational research is conducted is characterized by several, possibly conflicting purposes. There are the needs of professional education for school service, teacher training, curriculum development, etc., and there are the norms and values of research which seem to be affected adversely by the practice-orientation of professional education. We found that a large proportion of authors had practice-oriented socialization, that a small percentage had graduate research experience, that relatively few had specialized careers in research, and that a large proportion of authors preferred some activity other than research. These characteristics reflect professional goals that compete with research for attention. It is entirely possible that authors who contribute to the realization of practice-oriented goals are rewarded, regardless of the quality of their research. Since we have no way of assessing the quality and quantity of service activities performed by these authors, we cannot determine whether something else is more likely to be rewarded in education than good research. This problem is beyond our scope here, but it certainly warrants further study. We can only suggest this as one of several interpretations of why authors of better research are no more likely to be rewarded than authors of mediocre or poor research.

---

[1] Cole, "Social Structure of Science."

Summary

In this chapter we illustrated how the reward system operates in educational research, and we considered several consequences of the way it operates. Finally, we offered two possible interpretations of why it operates the way it does. Specifically, we discovered that better research in education is no more likely to be rewarded than is mediocre research. By way of contrast, in the behavioral sciences better research is somewhat more likely to be rewarded. We found that rewarded authors are more likely to prefer research than unrewarded ones.

We considered the possibility that the system is rewarding productivity instead of quality, but found only moderate evidence to support that interpretation. We also suggested that in education the reward system may operate the way it does because of the highly diffuse and inefficient communication system. We presented data from Nelson's study documenting the diffuse patterns of communication among some of the authors in this sample, and considered other evidence about the nature of the communication system. As an additional explanation of why the reward system operates the way it does in education, we noted the existence of multiple goals, with the attendant likelihood that non-research activities compete with research for the rewards of the system.

CHAPTER X

SUMMARY AND CONCLUSIONS

The objective of this study has been to measure the quality of recent research on education and to analyse the conditions that are related to better research. In theoretical terms the problem is one of specifying the processes which explain performance of the research role. We have examined the backgrounds, role socialization, and organizational settings of individual authors, and we have considered how the reward system operates within their research tradition. In this chapter we will briefly summarize the major findings of our study.

Before exploring the conditions associated with variation in research quality, we examined the nature of quality in some detail. We defined good research as that which contributes to theory, contributes to educational practice, and uses sound research methods or contributes to the development of new research methods. We then discussed the strengths and limitations of various measures of research including citation rates, evaluation forms, productivity, honorific awards, peer evaluations, journal characteristics, and self evaluations. In this study we decided that using a panel of experts to rate the quality of specific research papers was the most appropriate method for several reasons. First, having a panel of judges rate research papers does not require the assumptions that a method such as citation rates does.

Specifically, in view of evidence suggesting that even the best research in education is not always visible, we did not want to make the assumption of equal visibility that we feel is required by the use of citation rates. Second, focusing on a particular paper enabled us to relate the conditions under which a paper was written to its rated quality. Similarly, this method permitted us to explore rewards since the time a particular paper was written.

The chief requirements of expert evaluations of research quality are agreement on what the standards to assess research quality should be and agreement among the judges on how a paper should be rated on each of the three dimensions of quality. We discussed these conditions in Chapters III and IV. There we saw that three other efforts to rate research papers agreed upon the use of an overall, or global, rating form. This agreement, plus its greater ease of administration was our rationale for using a global form rather than an evaluation form comprised of numerous individual criteria. We used a global rating of three dimensions of quality--contribution to theory, contribution to practice, and use of research methods. When the three dimensions were averaged, the average deviation from the mean over the nine papers that were rated in common was .62, which is less than one adjacent category on the five point rating scale. In the analysis presented in this volume we virtually always collapsed the five point scale into three categories--above average, average, and poor. As a result, the agreement among judges within this broader category is even greater. The extent of agreement among judges gives confidence in the reliability of the rating form used in this study.

What do these ratings of research quality indicate about the nature of educational research? The distribution of ratings in the sample, presented in Chapter V, indicates that the quality of educational research is generally quite low. When ratings of the three dimensions are averaged, 14 per cent of the papers were rated "Incompetent," compared to only 5 per cent which were rated "Outstanding," on a five-point scale. On the compressed three-point scale, 39 per cent of the papers were rated "Below Average" or "Incompetent" compared to 30 per cent which were rated "Above Average" or "Outstanding." These findings confirm our impression, and that of many other observers of educational research, that research on education is generally quite poor.

A number of interpretations of why educational research is poor have been offered, and we considered them in this study. For example, some argue that the undergraduate performance of recruits entering educational research is crucial for better research. Some emphasize the importance of research apprenticeships. Some stress that school teaching should be avoided in preparing for a research career. Others mention that the service orientation of schools of education is detrimental for research. Still others have noted the lack of specialization in the careers of reading researchers and indicate that this affects research adversely.

None of the above studies measured the quality of educational research, however. Instead, they either had an impressionistic feeling that the quality was generally low and they offered explanations of this perceived condition; or, they measured something other than quality and assumed that it was related to better research. Wilder measured research involvement which we have seen is related to quality. Buswell and Worthen, however, measured productivity which we found is unrelated to better research in education.

In this study we have considered the above interpretations as hypotheses to test, but our basic approach has not been one of testing single hypotheses. We view quality as the outcome of a series of interdependent processes, and therefore we have analysed the interaction of recruitment, role socialization, work settings, larger research context (meaning the reward and communication systems) in relation to better research.

In Chapter VI we described and measured five key elements of role socialization--type of doctorate, practice-oriented socialization, graduate research experience, career research experience, and research orientation. After considering their individual relationships to research quality, we formulated a model of the interrelationships between these elements and quality. In the model we postulated that an author's socialization tended to direct him either toward the practice-oriented role of professional educator or toward the role of researcher, with major consequences for the quality of his research. Thus, education doctorates are more likely to have practice-oriented socialization, less likely to have graduate and career research experience and less likely to become research oriented, all of which has a cumulative negative effect on the quality of their research. In contrast, behavioral science doctorates are less likely to receive practice-oriented socialization, more likely to have graduate and career research experience and hence more likely to become research oriented, which has a positive cumulative effect on quality.

We returned to background in Chapter VII, and analyzed recruitment in relation to relevant aspects of role socialization. Sex (being female) and religion (being Jewish or areligious) were strongly related to doing better research even when type of doctorate, amount of practice-oriented socialization, research experience and research orientation were held constant. We found that the relationship between being female, Jewish, or areligious and better research partially disappeared when intellectual orientation (as measured by curiosity about substantive questions in a behavioral science field) was held constant. Undergraduate performance, an explanation of research quality offered by many observers of educational research, was found to have only a small positive relationship to better research, and it was unrelated to type of doctorate, which means that behavioral scientists are not doing better research because they were better undergraduate performers.

In Chapter VIII we shifted our focus from the authors' background and role socialization to the organizational setting in which they wrote their research papers. We considered four organizational settings which were identified by previous observers of educational research as possibly affecting quality. Perhaps most surprisingly in this chapter we learned that the negative association between working in a school of education and quality disappeared when type of doctorate was held constant. Thus, behavioral scientists in schools of education did just as good research as behavioral scientists in behavioral science departments. It was the greater tendency of authors in schools of education to have received doctorates from schools of education that seemed to explain why they were less likely to do good research. We learned also that being a university research center staff member is related to better research among research-oriented authors, although not among those with other orientations. Furthermore, the small positive relationship between departmental prestige and better research disappeared when working in a university research center was held constant.

In Chapter IX, we described and analyzed the operation of the reward and communication sub-systems in educational research. We found evidence that the reward system in education is no more likely to reward the authors of better research papers than the authors of mediocre or poor papers, whereas in the behavioral sciences better research is somewhat more likely to be rewarded than poorer research. We noted that the reward system serves to enhance authors' preference for research as a professional activity. Therefore, it is very likely that good researchers who have not been rewarded will lose interest in doing research. This process, in combination with the likely exposure of authors to practice-oriented socialization and the unlikely chance for considerable graduate or career research experience reduces the probability that authors will be research oriented, and thereby has serious negative consequences for the quality of research produced in education.

In Chapter IX we also tried to explain why the reward system in educational research operates the way it does. We reviewed the evidence for the "publish or perish" interpretation of how the reward system operates in educational research. While we found a stronger relationship between productivity and rewards than between quality and rewards, the evidence was not so consistently strong that we were convinced that frequent publication completely explains why the reward system operates the way it does in education. Therefore, we considered two other interpretations, namely the diffuse nature of the communication system and the existence of multiple goals in educational research, and noted that evidence exists to support both of these interpretations, at least on a tentative basis. It is plausible, of course, that all three of these interpretations multiply determine the operation of the reward system.

In general terms we have hypothesized in this study that performance of the research role is affected by a series of interdependent processes. Thus, we have adopted what might be called a "social systems" approach to the question of research quality. This general orientation has made our analytic task a difficult one to be sure; but we feel that the merit of this strategy has been borne out by our findings. Specifically, we have found that background characteristics, role socialization, structural settings, and the reward structure are interrelated and affect the quality of research either directly or through intervening processes.

We first formulated a model of how the elements of role socialization are related to each other and to better research. After testing this model and modifying it slightly, we proceeded to use it as the basis for further analysis. In particular, we considered how the background of recruits affected relationships between role socialization and quality, and examined the way socialization interacted with research contexts to produce variations in research quality. By using this model as our starting point, we supplemented our theory of role socialization with propositions about the effects of selective recruitment and structural settings.

We conclude by describing an "ideal" researcher and an "ideal" system in which to work, in terms of our findings. The most important individual characteristic of the "ideal" researcher is his intellectual orientation (as measured by "curiosity about substantive questions in a behavioral science field"). In addition, the "ideal" researcher has a series of experiences that facilitate socialization into the research role, specifically receiving a doctorate in the behavioral sciences rather than in education, having graduate and career research experience, and not having practice-oriented socialization (that is, not having elementary or secondary school teaching experience, having a Ph.D. instead of an Ed.D., and having taken no professional educational courses).

An important outcome of the above experiences is becoming research oriented. Research orientation seems to be a pivotal concept for understanding the processes related to better research. Prior research experience is highly related to research orientation, as is the lack of practice-oriented socialization. Research orientation, in turn, is highly related to better research. The "ideal" researcher, then, does better research by contributing to theory with new conceptualizations or hypothesized interrelationships; by contributing to educational practice (at least potentially); and by using sound methods. In short, the "ideal" researcher develops original insights based on sound information and inferences.

An "ideal" system in which to do research has the following characteristics:

(1) Patterns of recruiting intellectually oriented people;

(2) Patterns of socializing recruits to the research role, including:

- (a) the existence of opportunities for graduate research experience,
- (b) the existence of opportunities for career research experience,
- (c) the lack of practice-oriented socialization in the preparation of researchers.

(3) The existence of structures that facilitate research, e.g., university research centers and specialized agencies.

(4) Standards for evaluating the quality of research, and the uniform and universalistic[1] application of those standards.

(5) A system that rewards better research.

(6) A shared system of communication--consisting of a core of journals to which most researchers submit their papers and of common

---

[1] The importance of universalistic application of standards is stressed by Cole in "The Social Structure of Science."

professional association membership--to insure visibility of better work and to facilitate efficient operation of the reward system.

(7) Career specialization, so that non-research activities do not compete for the same time, attention, resources, and rewards as research activities.

All of the above social structural characteristics may be considered aspects of a single summarizing concept that might be called the institutionalization of research. The more of each of these characteristics in a field, that is the higher the level of institutionalization, the better the research produced in that field. Whether this proposition is applicable to all fields is a question that needs further study.

GLOSSARY

Career Research Experience refers to the extent to which an author has done research during his career. We formed an index of career research experience from the following three indicators: author spent more than 50 per cent of his time (when he wrote the research paper which was rated) doing research, author had a period of six months or more during his career when research was his primary activity, and author was a staff member of a research center at the time he wrote the paper. If an author had two or three of these experiences, his career research experience was considered high. If he had only one, it was considered medium, and if he had none, it was considered low. This index was very highly related to the average rank of research throughout an author's career. The index is introduced on page 170. See also Appendix VI-4 for an intercorrelation matrix of these items and for more discussion of how the index was constructed and validated.

Departmental Prestige refers to the reputation of an academic department. The prestige of education departments was based on the opinions of Deans of Schools of Education as reported by Sieber (in The Organization of Educational Research). Deans named 22 schools of education as ones that produced better research. The prestige of psychology and sociology graduate departments was based on

Cartter's reputational measures of those departments. See Appendix VI-3 for a more detailed description of how the index was formed.

Graduate Research Experience means the experience an author had actually doing research or helping someone else do it during graduate school, independently of his formal course work. We asked authors, "While in graduate school did you ever work as a research assistant?" Work as a research assistant is assumed to mean either that an author helped a university professor with his research or that he assisted on a research project that may or may not have been directed by a faculty member. We also asked, "As a graduate student did you ever work in a research center or bureau within the university?" Cross-tabulating the two questions resulted in three patterns of graduate research experience: experience both as a research assistant and in a research center, work as a research assistant but not in a research center, and neither. (Only 24 authors worked in centers but not as research assistants.) The term graduate research experience is introduced on page 167.

Institutionalization of research is a general summarizing concept that refers to patterns of recruitment, role socialization, career specialization, research settings, research standards, communications, and reward allocation. The term appears first on page 78.

Index of Having Been Rewarded: Authors were classified as having been rewarded if they had one or more of the following: if they published an average of more than two papers per year since their first publication, if they applied for and received a research grant, if they were promoted in academic rank between the time they wrote the paper and the time they answered the questionnaire, or if they moved from a less prestigious to a more prestigious department during that time period. See Appendix IX-4 for a more detailed description of how this index was constructed. The term first appears on page 288.

Intellectual Orientation refers to being curious about substantive questions in a behavioral science field. We measured it with the following questionnaire item: "Of course, many factors motivate individuals to do research. Regarding your usual research motives, how would you distribute 10 points among the following (so the total equals 10)?" The more points authors gave "curiosity about substantive questions in a behavioral science field," the greater their intellectual orientation. The term first appears on page 207.

Practice-Oriented Socialization means that an author had either graduate training or career experience that prepared him more for the role of educational practitioner rather than for the role of researcher. We formed an index of practice-oriented socialization out of the following items: having one or more years of elementary or secondary school teaching experience, taking one or more professional education courses, and earning an Ed.D. rather than a Ph.D. If an author had none of these experiences he was considered to have less practice-oriented socialization. If he had one, he was considered to have some such socialization. If he had two or three, he was considered to have more practice-oriented socialization. See Appendix VI-2 for a matrix of correlations among these items and for more description of this index. The term appears first on page 165.

Productivity refers to the quantity of research publications an author has produced. We measured this by computing an average annual number of research articles an author had published since his first publication, based on his own self report. We found that the number of books or monographs was related to other variables, such as quality, in the same way that the number of articles was, so we used only the average annual number of research articles. The term appears first on page 271. See Appendix IX-1 for more description of how the index of productivity was computed.

Research Quality refers to the assessed worth of particular research papers. A national panel of 39 judges rated 390 research papers on three dimensions of quality: their substantive contribution to theory, their substantive contribution to any field(s) of educational practice, and their utilization of (or contribution to) research methods. The judges rated each dimension on a five point scale consisting of the following categories:

(I) Ranks with the best empirical research studies known to me; on a par with the top 5 or 10 per cent in this respect.

(II) Better than average, though not "outstanding" in this respect.

(III) Run of the mill in this respect; neither better nor poorer than the bulk of research that I have seen.

(IV) Not up to "average" standards; "less-than-mediocre" in this respect (although not altogether lacking).

(V) Incompetent in this respect; among the poorest examples of "research" that I have encountered.

In most of our analysis we combined the "best" and "better than average" categories into one group of better research papers. This measure appears first in Chapter IV.

The Reward System refers to the process of allocating rewards in a field. We measured four processes of reward allocation: publishing frequently as indicated by the average annual number of research papers published by an author; receiving a research grant; being promoted in academic rank or moving from a less prestigious to a more prestigious academic department between the time the paper was written and the time the questionnaire was returned (a period of about five years). The term appears first on page 270.

(Role) Orientation means the positive valence an author feels toward a particular professional activity. We measured this by asking: "Please rank three of the following activities in terms of your personal preference (write 1 next to the job you would like to do most, 2 next to your second choice, and 3 next to your third choice)." Authors who ranked empirical research or research related activities (e.g., research administration) first were considered to be research oriented. We speak of them either as research oriented authors or as authors who prefer research. The term appears first on page 179.

Role Socialization is a concept comprised of type of doctorate, practice-oriented socialization, graduate research experience, career research experience, and (role) orientation. It appears first on page 153.

Type of Doctorate refers to whether the author received his doctorate from a school or department of education or from some other department. We measured type of doctorate by asking, "Was degree in education? Yes No." Almost all who replied "yes" majored in education or in an education field (such as music education, science education, etc.). We refer to these respondents as education doctorates. Since 90 per cent of the authors who did not receive their doctorates in schools of education majored in one of the behavioral sciences, we refer to those respondents as behavioral scientists. The term type of doctorate appears first on page 154.

Undergraduate Performance refers to an author's scholastic achievement as an undergraduate. We measured it as follows: first, undergraduate school selectivity was measured by using Berelson's ranking of colleges and universities, Cass and Birnbaum's evaluation of the admissions selectivity of colleges and universities, and an index of school resources developed at the Bureau of Applied Social Research. If a college or university was classified as high on any one of these three measures, it was considered selective. Second, we asked authors if they had won any academic honors as an undergraduate. If an author earned one or more academic honors in a selective undergraduate college or university, he was classified as being a high performer; if not, as being an average performer.

SELECTED BIBLIOGRAPHY

Delineating the Population

"America's Education Press." New York: Educational Press Association of America, Yearbook 29, 1966.

Bargar, Robert; Guba, Egon; and Okorodudu, Corahann. "Development of a National Register of Educational Researchers." U.S.O.E. Cooperative Research No. E-014. Columbus, Ohio: Ohio State University Research Foundation, December 1965.

Bibliographic Index. New York: H. W. Wilson Co., 1967, 1968.

Blaugh, M. *Economics of Education: A Selected Annotated Bibliography*. New York and Oxford: Pergamon Press, 1966.

Directory: Newspapers and Periodicals 1968. Philadelphia, Pa.: N. W. Ayer and Sons, 1968.

Economic Abstracts. Vol. LV-1, No. 21, February 1956. New York: New York University, 1956.

Education Index. New York: H. W. Wilson Co., Vol. 40, No. 1, 4.

Herner, Sau; Griffith, Janet D.; and Herner, Mary. "Study of Periodicals and Serials in Education." Final Report, Project No. BR 7-9003, U.S. Office of Education, Department of Health, Education, and Welfare. Washington, D.C.: Herner and Company, 1968.

Index of Economic Journals. Vol. VII, 1964-65. Prepared under the auspices of the American Economic Association. Homewood, Ill.: Richard D. Irwin, Inc., 1967.

International Bibliography of the Social Sciences: Economics, 1966. Vol. XV. Chicago: Aldine. London: Tavistock, 1966.

International Bibliography of the Social Sciences: Sociology. Chicago: Aldine. London: Tavistock, 1966.

International Bibliography of the Social Sciences: Social and Cultural Anthropology, Vol. XII. Chicago: Aldine. London: Tavistock. Prepared under the auspices of the International Union of Anthropological and Ethnological Sciences by ICSSD, with the support of UNESCO, 1967.

International Bibliography of the Social Sciences: Political Science, Vol. XV. Chicago: Aldine; London: Tavistock, 1966.

International Political Science Abstracts, Vol. XVII, No. 1, October-December 1966. Oxford: Basil Blackwell, 1967.

Psychological Abstracts. Washington, D.C.: American Psychological Association, 1967.

Publication Manual of the American Psychological Association. Washington, D.C.: American Psychological Association, 1967 revision.

Sociological Abstracts. New York: International Sociological Association, Eastern Sociological Society, and Midwest Sociological Society. Vol. 13, Nos. 1-7, covering 1965, c. 1964-67.

General Theoretical Orientations

Blau, Peter. *The Dynamics of Bureaucracy*. Chicago: University of Chicago Press, 1955.

Broom, Leonard. "Social Differentiation and Stratification." *Sociology Today*. Edited by Robert K. Merton, Leonard Broom, and Cottrell, Jr. New York: Harper Torchbooks, 1959, Vol. II.

Clark, Terry N. "Institutionalization of Innovations in Higher Education: Four Conceptual Models." *Administrative Science Quarterly*, XIII, No. 1 (June, 1968), 1-25.

Durkheim, Emile. *Education and Sociology*. Translated by Sherwood D. Fox. Glencoe, Ill.: Free Press, 1956.

Epstein, Cynthia F. "Encountering the Male Establishment: Sex-Status Limits on Women's Careers in the Professions." *American Journal of Sociology*, LXXV, No. 6 (May, 1970), 965-82.

________. *Woman's Place: Options and Limits in Professional Careers*. Berkeley: University of California Press, 1971.

Ericson, Erik. "Inner and Outer Space: Reflections on Womanhood." Daedalus, XCIII, No. 2 (Apring, 1964), 582-606.

Goode, William J. "Community Within a Community: The Professions." *American Sociological Review*, XXII (April, 1957), 194-200.

________. "Encroachment, Charlatanism, and the Emerging Professions: Psychology, Sociology, and Medicine." *American Sociological Review*, XXV, No. 6 (December, 1960), 902-11.

Gross, Neal. "The Sociology of Education." Sociology Today. Edited by Robert K. Merton, Leonard Broom, and Leonard S. Cottrell, Jr. New York: Basic Books, 1959.

________; Mason, Ward S.; and McEachern, Alexander W. Explorations in Role Analysis: Studies of the School Superintendency Role. New York: John Wiley & Sons, Inc., 1958.

Katz, Daniel, and Kahn, Robert L. The Social Psychology of Organizations. New York: John Wiley, 1966.

Kuhn, Thomas S. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.

Malewski, Andrzej. "The Degree of Status Incongruence and Its Effects." Class, Status, and Power. Edited by Reinhard Bendix and Seymour Lipset. 2nd edition. New York: Free Press, 1966.

Merton, Robert K. Social Theory and Social Structure. Revised and enlarged edition. New York: Free Press of Glencoe, 1957.

Rosen, Bernard C., and Bates, Alan P. "The Structure of Socialization in Graduate School." Sociological Inquiry, XXXVII (Winter, 1967), 71-84.

Sherlock, Basil J., and Morris, Richard T. "The Evolution of the Professional: A Paradigm." Sociological Inquiry, XXXVII (Winter, 1967), 27-46.

Simpson, Ida Harper. "Patterns of Socialization into Professions: The Case of Student Nurses." Sociological Inquiry, XXXVII (Winter, 1967), 47-54.

Smelser, Neil J. "Sociology and the Other Social Sciences." The Uses of Sociology. Edited by Paul F. Lazarsfeld, William H. Sewell, and Harold Wilensky. New York: Basic Books, Inc., 1967.

Snoek, J. Diedrick. "Role Strain in Diversified Role Sets." American Journal of Sociology, LXXI, No. 4 (1966), 363-72.

Tolman, Edward C. "A Psychological Model." Toward a General Theory of Action. Edited by Talcott Parsons and Edward A. Shils. New York: Harper Torchbooks, 1951.

Weber, Max. The Protestant Ethic and the Spirit of Capitalism. Translated by Talcott Parsons, 1930. New York: Charles Scribner's Sons, 1958.

Wheeler, Stanton. "The Structure of Formally Organized Socialization Settings." Socialization After Childhood. Edited by Orville G. Brim and Stanton Wheeler. New York: Wiley, 1966.

Zurcher, Louis A., Jr. "The Naval Recruit Training Center: A Study of Role Assimilation in a Total Institution." Sociological Inquiry, XXVII (Winter, 1967), 85-98.

Measuring Research Output

Bartos, Bruce B. "A Review of Instruments Developed to be Used in the Evaluation of the Adequacy of Reported Research." Bloomington, Indiana: Phi Delta Kappa, 1969. (Mimeographed.)

Chase, Janet M. "Normative Criteria for Scientific Publication." The American Sociologist, V, No. 3 (August, 1970), 262-65.

Eiduson, Bernice T. "Productivity Rate in Research Scientists." American Scientist, LIV, No. 1 (1966), 57-63.

Gephart, William J. "Development of an Instrument for Evaluating Reports of Educational Research." Unpublished Ph.D. dissertation, The Ohio State University, 1965.

Nasatir, David, and Elish, David. "Project Memorandum No. 5" of The Organization of Educational Research. New York: Bureau of Applied Social Research, Columbia University, 1965. (Mimeographed.)

Orr, Richard H., et al. Peer Group Judgment of Scientific Merit. Philadelphia: Institute for Advancement of Medical Communication, in process.

Persell, Caroline Hodges. "Measuring Educational Research Quality and Its Correlates." Unpublished master's essay, Columbia University, 1967.

Sieber, Sam D. "Detailed Criteria of Evaluation." Rev. ed. New York: Bureau of Applied Social Research, Columbia University, 1966. (Mimeographed.)

Shaw, Byron T. "The Use of Quality and Quantity of Publication as Criteria for Evaluating Scientists." Miscellaneous Publication No. 1041. Washington, D.C.: Agricultural Research Service, U.S. Department of Agriculture, January 1967.

Tripodi, Tony; Fellin, Phillip; and Meyer, Henry J. The Assessment of Social Research: Guidelines for the Use of Research in Social Work and Social Science. Itasca, Ill.: F. E. Peacock, 1969.

Wandt, Edwin. A Cross-Section of Educational Research. New York: David McKay Company, Inc., c. 1965.

________; Adams, Georgia S.; Collett, Dorothy M.; Michael, William B.; Ryans, David G.; and Shay, Carleton B. "An Evaluation of Educational Research Published in Journals." Report of the Committee on Evaluation of Research, American Educational Research Association, February 1967. (Mimeographed.)

Research Methods

Barton, Allen H., and Lazarsfeld, Paul F. "Some Functions of Qualitative Analysis in Social Research." Bobbs-Merrill Reprint Series in the Social Sciences, S-336. Indianapolis, Indiana: Bobbs-Merrill Co., Inc.

Blalock, Hubert M., and Blalock, Ann, eds. Methodology in Social Research. New York: McGraw-Hill, 1968.

Borgatta, Edgar F., ed. Sociological Methodology, 1969. San Francisco: Jossey-Bass Inc., 1969.

Campbell, Donald, and Stanley, Julian. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Co., 1963.

Coleman, James S.; Katz, Elihu; and Menzel, Herbert. "The Diffusion of an Innovation among Physicians." Sociometry, XX (December, 1957), 253-70.

________; Katz, Elihu; and Menzel, Herbert. Medical Innovation: A Diffusion Study. Indianapolis, Indiana: Bobbs-Merrill Co. Inc., 1966.

________; Menzel, Herbert; and Katz, Elihu. "Social Processes in Physicians' Adoption of a New Drug." Journal of Chronic Diseases, IX, No. 1 (January, 1959), 1-19.

Etzioni, Amitai. A Comparative Analysis of Complex Organizations. New York: Free Press of Glencoe, 1961.

Glock, Charles Y. Survey Research in the Social Sciences. New York: Russell Sage Foundation, 1967.

Heise, David R. "Problems in Path Analysis and Causal Inference." Sociological Methodology. Edited by E. F. Borgatta. San Francisco: Jossey-Bass, 1968.

Kadushin, Charles. "Individual Decisions to Undertake Psychotherapy." Administrative Science Quarterly, III, No. 3 (December, 1958), 379-411.

________. "Reason Analysis." International Encyclopedia of the Social Sciences, XIII, 338-43. New York: Macmillan Company and Free Press, 1968.

Katz, Elihu, and Lazarsfeld, Paul F. Personal Influence. Glencoe, Ill.: Free Press, 1955.

Lazarsfeld, Paul F. "The Art of Asking Why." National Marketing Review, I, No. 1 (Summer, 1935). Also in Public Opinion and Propaganda. Edited by Daniel Katz, et al. New York: Dryden Press, 1954, pp. 675-86.

________. "Evidence and Inference in Social Research." Daedalus, LXXXVII, No. 4 (1958). Also Bobbs-Merrill Reprint Series in the Social Sciences, S-441. Indianapolis, Indiana: Bobbs-Merrill Co. Inc.

________. "The Statistical Analysis of Reasons as a Research Operation." Sociometry, V (1942), 29-47.

Robinson, W. S. "The Statistical Measurement of Agreement." American Sociological Review, XXII, No. 1 (February, 1967), 17-25.

Rossi, Peter H. Why Families Move: A Study in the Social Psychology of Urban Residential Mobility. Glencoe, Ill.: Free Press, 1955.

Sewell, William H.; Haller, Archibald O.; and Portes, Alejandro. "The Educational and Early Occupational Attainment Process." American Sociological Review, XXXIV, No. 1 (February, 1969), 82-92.

Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Co., Inc., 1956.

Sills, David L. "On the Art of Asking 'Why Not'?" Report of the Proceedings, Fourth All India Conference on Family Planning, Hyderabad, January 29-February 3, 1961. Bonbay, India: Family Planning Association of India.

Zeisel, Hans. Say It with Figures. 4th ed. New York: Harper & Brothers, 1957.

Zetterberg, Hans L. On Theory and Verification in Sociology. Stockholm: Almqvist and Wiksell; New York: Tressler Press, 2nd ed., 1957.

Research Training, Careers, and Contexts

American Institutes for Research. Report to the Board of Directors. October 26, 1964. (Xeroxed.)

Astin, Helen S. The Woman Doctorate in America. New York: Russell Sage Foundation, 1969.

Bargar, Robert R.; Okorodudu, Corahann P.; Dworkin, Edward P. "Investigation of Factors Influencing the Training of Educational Researchers." Final Report, Project No. 3191, Contract No. OE-6-10-146. Columbus, Ohio: Ohio State University Research Foundation, May 1970.

Berelson, Bernard. Graduate Education in the United States. New York: McGraw-Hill Book Co., Inc., 1960.

Bernard, Jessie. Academic Women. New York: Meridian Books, World Publishing Co., 1964.

Bloom, Benjamin S. "Twenty-Five Years of Educational Research." Educational Research and Contemporary Social Problems. Report of the Sixth Annual Convocation on Educational Research, October 18-19, 1965. Sponsored by the University of the State of New York and the Educational Research Association of New York State.

Brown, L. D. Doctoral Graduates in Education: An Inquiry into Their Motives, Aspirations, and Perception of the Program. Cooperative Research Project No. S-240. Washington, D.C.: Office of Education, U.S. Dept. of Health, Education, and Welfare, 1966.

Brown, Leiba. "Educational Research and the Liberal Arts." New York: Bureau of Applied Social Research, Columbia University, May 1966.

Buswell, Guy T.; McConnell, T. R.; Heiss, Ann M.; and Knoell, Dorothy M. "Training for Educational Research." Cooperative Research Project No. 51074. Berkeley: University of California, 1966.

Cartter, Allan M. An Assessment of Quality in Graduate Education. Washington, D.C.: American Council on Education, 1966.

Cass, James, and Birnbaum, Max. Comparative Guide to American Colleges. New York: Harper and Row, 1965.

Charters, W. W. "The Social Background of Teaching." Handbook of Research on Teaching. Edited by N. L. Gage. Chicago: Rand McNally, 1963, pp. 715-813.

Clark, David L., and Hopkins, John E. "A Report on Educational Research, Development, and Diffusion Manpower, 1964-1974." Bloomington, Indiana: Indiana University Research Foundation, 1969.

________, and Worthen, Blaine R., eds. Preparing Research Personnel for Education. Bloomington, Indiana: Phi Delta Kappa and American Educational Research Association, 1967.

Clark, Kenneth E. America's Psychologists: A Survey of a Growing Profession. Washington, D.C.: American Psychological Association, 1957.

Coler, Myron B., ed. Essays on Creativity in the Sciences. New York: New York University Press, 1963.

Corwin, Ronald G., and Seider, Maynard. "Patterns of Educational Research: Reflections on Some General Issues." Paper prepared for a colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968. (Mimeographed.)

Cronbach, Lee J., and Suppes, Patrick, eds. Research for Tomorrow's Schools. New York: Macmillan Co., 1969.

David, Deborah. "Career Patterns and Values: A Study of Men and Women in Science and Engineering." Unpublished Ph.D. dissertation, Columbia University, 1971.

"Educational Research and Development in the United States." OE-12049. Washington. D.C.: U.S. Dept. of Health, Education, and Welfare, Office of Education, December 1969.

Eidell, Terry L., and Kitchel, Joanne M. Knowledge Production and Utilization in Educational Administration. Published jointly by University Council for Educational Administration, Columbus, Ohio, and Center for the Advanced Study of Educational Administration, University of Oregon, 1968.

Etzioni, Amitai, ed. The Semi-Professions and Their Organization: Teachers, Nurses, Social Workers. New York: Free Press, 1969.

Fattu, N. A. A Survey of Educational Research and An Appraisal by Scientists from Other Fields. Bloomington, Indiana: School of Education, Indiana University, 1967.

Foshay, Arthur W. "Issues and Dilemmas in Nurturing the Educational Research in an Organizational Setting." The Training and Nurture of Educational Researchers. Edited by Egon Guba and Stanley Elam. Bloomington, Indiana: Phi Delta Kappa, 1965.

Friedman, Robert A. "The Role of Technology and Occupation in Institutional Change: A Historical Case Study of American Public Education." Paper presented at the 65th Annual Meeting of the American Sociological Association, September 1970, Washington, D.C.

Glass, Gene V. *Some Observations on Training Educational Researchers.* Research Paper No. 22. Boulder: Laboratory of Educational Research, University of Colorado, November 22, 1968.

Guba, Egon, and Elam, Stanley. *The Training and Nurture of Educational Researchers.* Bloomington, Indiana: Phi Delta Kappa, Inc., 1965.

Hagstrom, Warren O. "Educational Researchers, Social Scientists, and School Professionals." Paper prepared for a Colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968. (Dittoed.)

Kendall, Patricia. "Impact of Training Programs on the Young Physician's Attitudes and Experiences." *Journal of American Medical Association*, June 24, 1961, pp. 992-97.

________, and Merton, Robert K. "Medical Education as Social Process." New York: Bureau of Applied Social Research, Columbia University.

________. "The Relationship between Medical Educators and Medical Practitioners." *Annals of the New York Academy of Sciences*, CXXVIII, Article 2 (September 27, 1965), 568-76.

Knapp, Robert H., and Goodrich, H. B. *Origins of American Scientists.* Chicago: University of Chicago Press, 1952.

Lazarsfeld, Paul F., and Sieber, Sam D. *Organizing Educational Research.* Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Levien, Roger E. "National Institute of Education: Preliminary Plan for the Proposed Institute." Washington, D.C. RAND Corporation, December 15, 1970.

Lortie, Dan C. "The Balance of Control and Autonomy in Elementary School Teaching." *The Semi-Professions and Their Organization.* Edited by Amitai Etzioni. New York: Free Press, 1969, pp. 1-53.

Merz, Louise E. "The Graduate School as a Socializing Agency: A Pilot Study of Sociological Aspects of Graduate Training in the Physical Sciences." Unpublished doctoral dissertation, Cornell University, 1961.

Millikan, Nancy M. "The Development of Professional Personnel in Educational Research." USOE Cooperative Research Project No. S-487-64. Unpublished doctoral dissertation, Teachers Dollege, Columbia University, 1967.

Nelson, Carnot E. (with the assistance of William Garvey and Nan Lin). "The Postmeeting Journal Dissemination of Material Presented at the 1968 American Educational Research Association Annual Meeting." Paper presented at the annual meeting of the American Educational Research Association, Minneapolic, Minnesota, March 1970. (Mimeographed.)

Persell, Caroline Hodges. "American Educational Research Association (AERA) Membership Data." New York: Bureau of Applied Social Research, Columbia University, 1970. (Xeroxed.)

________. "An Evaluation of a Strategy for Bringing Social Science to Bear on Organizational Decision-Making." New York: Bureau of Applied Social Research, Columbia University, April 1970.

Peterson, Osler L.; Andrews, Leon P.; Spain, Robert S.; and Greenberg, Bernard G. "An Analytical Study of North Carolina General Practice 1953-1954." The Journal of Medical Education, Part 2 (December, 1956).

Roaden, Arliss L. "Formalizing the Research Assistantship in Educational Research Training Programs." Paper prepared for a symposium, "The Research Assistantship as a Formal Component of Programs for Teaching Educational Researchers," conducted at the convention of the American Educational Research Association, Minneapolis, Minnesota, February 1970.

Roby, Pamela. "Structural and Internalized Barriers to Women in Higher Education." Washington, D.C.: Department of Sociology and Center for Manpower Policy Studies, George Washington University, 1970. (Mimeographed.)

Roe, Anne. "Women in Science." Personnel and Guidance Journal, XLIV (April, 1966), 8.

Rossi, Alice S. "Barriers to the Career Choice of Engineering, Medicine or Science among American Women." Women and the Scientific Professions. Edited by Jacquelyn A. Mattfeld and Carol G. Van Aken. Cambridge, Mass.: M.I.T. Press, 1965.

________. "Equality Between the Sexes: An Immodest Proposal." Daedalus, XCIII (Apring, 1964), 607-52.

________. "Women in Science: Why So Few?" Science, May 28, 1965, 1196-1202.

Sieber, Sam D. "Institutional Settings for Educational Research in the U.S." Educational Change in the U.S. Edited by Egon Guba. Bloomington, Indiana: Indiana University, 1967. Prepared for UNESCO.

________. "Observations on the Membership of the American Educational Research Association." New York: Bureau of Applied Social Research, Columbia University, 1966. (Mimeographed.)

________. The Organization of Educational Research. U.S.O.E. Cooperative Research Project No. 1974. New York: Bureau of Applied Social Research, Columbia University, 1966.

________. Reforming the University. Chicago: University of Chicago Press, forthcoming.

Simon, Rita; Clark, Shirley M.; and Galway, Kathleen. "The Woman Ph.D.: A Recent Profile." Social Problems, XV, No. 2 (Fall, 1967), 221-36.

Vollmer, Howard M. "A Preliminary Investigation and Analysis of the Role of Scientists in Research Organizations." Menlo Park, California: Stanford Research Institute, February 1962.

Wasserman, E. R., and Switzer, E. E. The Random House Guide to Graduate Study in the Arts and Sciences. New York: Vintage Books, 1967.

Wilder, David E. "The Reading Experts: A Case Study of the Failure to Institutionalize an Applied Science of Education." Unpublished Ph.D. dissertation, Columbia University, 1966. To be published by Teachers College, Columbia University, in press.

Wilson, Logan. The Academic Man: A Study in the Sociology of a Profession. New York: Octagon Press, 1964.

Worthen, Blaine R. "The Impact of Research Assistantship Experience on the Subsequent Development of Educational Researchers." Unpublished dissertation, The Ohio State University, 1968.

Sociology of Science

Barber, Bernard. Science and the Social Order. Glencoe, Ill.: Free Press, 1952.

________, and Hirsch, Walter, eds. The Sociology of Science. New York: Free Press, 1962.

Ben-David, Joseph. "Scientific Productivity and Academic Organization in 19th Century Medicine." Science and Society. Edited by Norman Kaplan. Chicago: Rand McNally, 1965, pp. 39-61.

________, and Collins, Randall. "Social Factors in the Origin of a New Science: The Case of Psychology." *American Sociological Review*, XXXI, No. 4 (August, 1966), 451-65.

Boalt, Gunnar. *The Sociology of Research*. Carbondale and Edwardsville: Southern Illinois University Press, 1969.

Cole, Jonathan R. "The Social Structure of Science: A Study of the Reward and Communications Systems of Modern Physics." Unpublished doctoral dissertation, Columbia University, 1969.

Cole, Stephen, and Cole, Jonathan R. "Scientific Output and Recognition: A Study in the Operation of the Reward System in Science." *American Sociological Review*, XXXII (June, 1967), 377-90.

________, and Cole, Jonathan R. "Visibility and Structural Bases of Awareness in Science." *American Sociological Review*, XXXIII, No. 3 (June, 1968), 397-412.

Collins, Randall. "Competition and Social Control in Science: An Essay in Theory Construction." *Sociology of Education*, XLI, No. 2 (Spring, 1968), 123-40.

Crane, Diana. "Scientists at Major and Minor Universities: A Study of Productivity and Recognition." *American Sociological Review*, XXX (October, 1965), 699-714..

________. "Social Structure in a Group of Scientists: A Test of the 'Invisible College' Hypothesis." *American Sociological Review*, XXXIV, No. 3 (June, 1969), 335-52.

________. "Social Class Origin and Academic Success: The Influence of Two Stratification Systems on Academic Careers." *Sociology of Education*, XLII, No. 1 (Winter, 1969), 1-17.

Garvey, William D.; Nelson, Carnot; and Lin, Nan. "A Preliminary Description of Scientific Information Exchange in Educational Research." Paper prepared for a Colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968. (Mimeographed.)

________, and Griffith, Belver C. "Scientific Communication as a Social System." *Science*, Vol. 167, Sept. 1, 1967, pp. 1011-16.

________; Lin, Nan; and Nelson, Carnot E. "Some Comparisons of Communication Activities in the Physical and Social Sciences." Baltimore: Johns Hopkins University, Center for Research in Scientific Communication, Department of Psychology, January 1970.

Glaser, Barney G. *Organizational Scientists: Their Professional Careers*. Indianapolis: Bobbs Merrill, 1964.

Hagstrom, Warren O. The Scientific Community. New York: Basic Books, 1965.

Innovations in Scientific Communication in Psychology. Washington, D.C.: American Psychological Association, 1966.

Kaplan, Norman. Science and Society. Chicago: Rand McNally, 1965.

Lin, Nan; Garvey, William D.; and Nelson, Carnot E. "A Study of the Communication Structure of Science." Baltimore: Johns Hopkins University, Center for Research in Scientific Communication, Department of Psychology, January 1970.

Lingwood, David A. "Interpersonal Communication, Scientific Productivity, and Invisible Colleges: Studies of Two Behavioral Science Research Areas." Paper prepared for a Colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968. (Xeroxed.)

Menzel, Herbert. "The Flow of Information among Scientists: Problems, Opportunities, and Research Questions." Prepared for the National Science Foundation. New York: Bureau of Applied Social Research, Columbia University, May 1958.

________. "Innovation, Integration, and Marginality: A Survey of Physicians." American Sociological Review, XXV, No. 5 (October, 1960), 704-13.

________. "Scientific Communication: Five Sociological Themes." American Psychologist, XXI (November, 1966), 999-1004.

________, and Katz, Elihu. "Social Relations and Innovation in the Medical Profession: The Epidemiology of a New Drug." The Public Opinion Quarterly, XIX, No. 4 (Winter, 1955-56), 337-52.

Merton, Robert K. "Basic Research and Potentials of Relevance." American Behavioral Scientist, VI, No. 9 (May, 1963), 86-90.

________. "The Matthew Effect in Science." Science, January 5, 1968, pp. 56-63.

Paisley, William J. "The Flow of (Behavioral) Science Information: A Review of the Research Literature." Institute for Communication Research, Stanford University, November 1965. 3rd publishing: December 1966.

________. "The Role of Invisible Colleges." Paper prepared for a Colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968. (Mimeographed.)

Price, Derek. *Little Science, Big Science*. New York: Columbia University Press, 1963.

Storer, Norman W. "The Hard Sciences and the Soft: Some Sociological Observations." *Bulletin of the Medical Library Association*, LV, No. 1 (January, 1967), 75-84.

————————. "The Organization and Differentiation of the Scientific Community: Basic Disciplines, Applied Research and Conjunctive Domains." Prepared for a Colloquium on "Improving the Social and Communication Mechanisms of Educational Research," sponsored by the American Educational Research Association, Washington, D.C., November 21-22, 1968. (Xeroxed.)

————————. "The Sociology of Science: Major Directions for Future Research." August, 1966. (Mimeographed.)

————————. "Science and Scientists in an Agricultural Research Organization: A Sociological Study." Unpublished doctoral dissertation, Cornell University, 1961.

————————. *The Social System of Science*. New York: Holt, Rinehart, and Winston, 1966.

Zuckerman, Harriet Anne. "Nobel Laureates in the United States: A Sociological Study of Scientific Collaboration." Unpublished doctoral dissertation, Columbia University, New York, 1965.

————————. "Nobel Laureates in Science: Patterns of Productivity, Collaboration, and Authorship." *American Sociological Review*, XXXII, No. 3 (June, 1967), 391-403.

————————. "The Sociology of the Nobel Prizes." *Scientific American*, CCXVII, No. 5 (November, 1967), 25-33.

————————. "Stratification in American Science." *Sociological Inquiry*, XL (Apring, 1970), 235-57.

————————, and Merton, Robert K. "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System." *Minerva*, IX, No. 1 (January, 1971), 66-100.

Miscellaneous

"Field Reader Catalogue." Office of Education, U.S. Dept. of Health, Education, and Welfare. Prepared by Autonetics Division of North American Rockwell Corp., Contract No. OEC-4-6-001305-1017. August, 1968.

Havelock, Ronald. "A Comparative Study on the Dissemination and Utilization of Scientific Knowledge." Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

McClelland, David C. "On the Psychodynamics of Creative Physical Scientists." Contemporary Approaches to Creative Thinking. Edited by H. E. Gruber, G. Terrall, and Mr. Wertheimer. New York: Atherton Press, 1962.

Nosow, Sigmund, and Form, William H., eds. Man, Work, and Society: A Reader in the Sociology of Occupations. New York: Basic Books, Inc., 1962.

Scriven, Michael. "The Philosophy of Science in Educational Research." Review of Educational Research, XXX (December, 1960).

Strodtbeck, Fred L. "Family Interaction, Values, and Achievement." Talent and Society. Edited by David C. McClelland, et al. New York: Van Nostrand, 1958.

APPENDIX XX-1

JOURNALS CONSIDERED BUT ELIMINATED WITHOUT SCANNING
(N=131)

	Reason for Elimination		
Name	COULDN'T FIND*	OUT OF DEFINITION OF POPULATION	OTHER
AEDS Journal	X		
Aerospace		X	
African Studies		X	
African Studies Bulletin		X	
Agricultural Education Magazine		X	
ALA Bulletin (American Library Association)		X	
American Artist		X	
American Economist		X	
American Forests		X	
American Institute of Planners Journal		X	
American Scientist		X	
American Statistician		X	
Anthropological Records of the University of California	X		
Archaeology		X	

---

*Before a journal was eliminated for this reason, the following libraries were checked: Bank Street College of Education library, Business library of Columbia University, Engineering library of Columbia, General library of Columbia, and Psychology library of Columbia, New York Public Library--42 Street and Donnell Education Library, New York University Education Library, New York University General Library, and Teachers College Library of Columbia University.

Name	Reason for Elimination		
	COULDN'T FIND	OUT OF DEFINITION OF POPULATION	OTHER
Art in America		X	
Art International		X	
Art Journal		X	
Art News		X	
Arts and Activities		X	
Audubon Magazine		X	
Automated Education Letter	X		
Aviation Week and Space Technology		X	
Behavior	X		
Biblical Archaeologist		X	
Bulletin of the Council for Research in Music Education	X		
Bulletin on Conservation Education		X	
Bulletin on International Education	X		
Business Automation		X	
Chemical and Engineering News		X	
Child and Family		X	
Chronicle of Higher Education			Not Empir'l.
College Business		X	
Contemporary Psychoanalysis		X	
Data Processing Magazine		X	
Data Processor		X	Not Empir'l.

Name	Reason for Elimination		
	COULDN'T FIND	OUT OF DEFINITION OF POPULATION	OTHER
Datamation		X	
Design Quarterly		X	
Discipline	X		
Econometrica		X	
Economic Development and Cultural Change		X	
Economic Record		X	
Education Age	X		
Education and Psychology Review			Review
Education Consultant	X		
Education Today			Newsletter (No Empir'l.)
Educator's Dispatch	X		
Educational and Psychological Interactions	X		
Federal Probation		X	
Film News		X	
Folia Linguistica		X	
Foreign Affairs		X	
General Linguistics	X		
Georgia Psychologist	X		
Glossa		X	
Graduate Research in Education and Related Disciplines	X		
Ideas Educational	X		

Name	Reason for Elimination		
	COULDN'T FIND	OUT OF DEFINITION OF POPULATION	OTHER
International Journal of Psychiatry		X	
International Journal of Sociometry and Sociatry		X	
Journal of African Languages		X	
Journal of American Oriental Society		X	
Journal of Engineering Education		X	
Journal of Health, Physical Education and Recreation (AAHPER)		X	
Journal of Home Economics		X	
Journal of Individual Psychology		X	
Journal of Industrial Arts Education		X	No Empir'l.
Journal of Industrial Teacher Education		X	No Empir'l.
Journal of Medical Education		X	
Journal of Outdoor Education		X	
Journal of True Education (Seventh Day Adventists)		X	
Journal of West African Languages		X	
Journalism Quarterly		X	
Kaiser Aluminum News		X	
Kroeber Anthropological Society Papers (Berkeley, Calif.)	X		

Name	**Reason for Elimination** COULDN'T FIND	OUT OF DEFINITION OF POPULATION	OTHER
Linguistics		X	
Mechanical Translation		X	
Michigan Psychologist		X	
Minnesota Studies in Vocational Rehabilitation	X	X	
Missouri Journal of Research in Music Education	X		
Modern Schoolman	X		
Monographs of the Society for Research in Child Development		X	
Museum News		X	
Nation's Business		X	
Nature Study		X	
Neurology		X	
North Central Association Quarterly			Regional Association Journal
Nursing Outlook		X	
Nursing Research		X	
Our Public Lands		X	
Outdoor Teacher		X	
Phi Kappa Phi Journal			No empir'l.
Phonetica		X	
Physics Today		X	
Political Affairs		X	
Practical Anthropology	X		
Product Information for Schools		X	
Product News		X	

Name	Reason for Elimination		
	COULDN'T FIND	OUT OF DEFINITION OF POPULATION	OTHER
Progressive Architecture		X	
Progressive Teacher	X		
Psychological Monographs			Monograph
Psychological Record		X	
Psychological Research Bulletin	X		
Psychological Researches	X		
Publication of the American Dialect Society		X	
The Quarterly (Buffalo)			Circulation is too small (under 1000)
Regional Science Association Papers and Proceedings		X	Proceedings not included in Pop.
Reports (University of Chicago)	X		
School Product News		X	
School Shop		X	
Science and Technology		X	
Scientific Research		X	
SEE (Self Enhancing Education)	X		
Sky and Telescope		X	
Slavic and East European Journal		X	
Social and Economic Studies			Foreign publication
SRIS Quarterly	X		
Studia Linguistica		X	

Name	Reason for Elimination		
	COULDN'T FIND	OUT OF DEFINITION OF POPULATION	OTHER
Studies and Research	X		
Studies in Higher Education	X		
Studies in Linguistics	X		
Studies on the Left		X	
Universities Quarterly			Foreign (London)
Videobriefs		X	
VISTA Volunteer		X	Not Empir'l.
Visual Communications Instructor		X	
Vocatio (Regional Vocational Education Publication)	X		
Wall Street Journal		X	
Western Economic Journal	X	X	
Wilson Library Bulletin		X	
Winnower	X		
Wisconsin Sociologist	X		
Word		X	

APPENDIX II-2

JOURNALS SCANNED BUT NOT YIELDING ARTICLES

N = 259

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
AAUP Bulletin		X		
AAUW Journal		X		
ACT Research Reports				Monograph
AGB Reports		X		
Adult Education (London)		X		Foreign
Adult Leadership		X		
Africa	X			
American Annals of the Deaf			X	
American Anthropologist	X			
American Biology Teacher		X		
American Documentation	X			
American Economic Review	X			
American Education		X		
American Historical Review	X			
American Journal of Archaeology	X			
American Journal of Economics and Sociology	X			
American Journal of Physics	X			
American Journal of Psychoanalysis	X			

2.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
American Journal of Psychology	X			
American Journal of Psychotherapy	X			
American Mathematical Monthly	X			
American Psychologist			X	
American Scholar		X		
American School Board Journal		X		
American School and University		X		
American Speech	X			
American Vocational Journal		X		
Annals, The (American Academy of Political and Social Science)		X		
Anthropological Linguistics	X			
Anthropological Quarterly	X			
Architectural Record			X	
Art Education		X		
ASHA (American Speech and Hearing Association			X	
AST Bulletin (Association for Student Teaching Bulletin)		X		
Atlantic Monthly		X		
Audio Visual Instruction		X		
Balance Sheet		X		
Behavioral Research and Therapy (England)				Foreign

3.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Behavioral Science			X	
Bioscience	X			
British Journal of Educational Psychology				Foreign
British Journal of Philosophy of Science				Foreign
Business Education Forum		X		
Canadian Journal of Linguistics				Foreign
Catalyst			X	
Catholic Education Review		X		
Catholic School Journal		X		
CBE Bulletin (Council for Basic Education)		X		
Central States Speech Journal			X	
Challenge				Stopped publication 8/19/67.
Child Study				Stopped publication in 1960.
Childhood Education		X		
Children		X		
Clearing House		X		
College Management		X		
College and Research Libraries			X	
College and University Business		X		
College and University Journal		X		

4.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Commentary		X		
Communications of the ACM	X			
Comparative Education (Oxford)				Foreign
Comparative Education Review				Review
Comparative Studies in Society and History (Hague)				Foreign
Computers and Automation			X	
Computers in the Humanities			X	
Contemporary Psychology			X	
Continuing Education for Adults				Out of Print June 1968
Crisis		X		
Current Anthropology (Utrect)				Foreign
Data Processing for Education		X		
Education Abstracts				Out of Print 1964
Education Digest		X		
Education and Training of the Mentally Retarded		X		
Education Reporter		X		Circulation less than 1000
Education Summary		X		
Education USA		X		

5.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Educational Administration Quarterly		X		
Educational Broadcasting Review		X		
Educational Forum		X		
Educational Horizons		X		
Educational Leadership		X		
Educational Research (England)				Foreign
Educational Research Service Circular (NEA)		X		
Educational Screen and Audio-Visual Guide		X		
Educational Technology		X		
Educational Theatre Journal		X		
Educational Theory		X		
Elementary English		X		
English Journal			X	
E.P.I.E. Forum		X		
Ethnology	X			
Eugenics Quarterly	X			
Exchange		X		
French Review		X		
Foundations of Language (Holland)				Foreign
German Quarterly		X		
Gifted Child Quarterly		X		

6.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Grade Teacher		X		
Harper's		X		
High School Journal		X		
Hispania		X		
Hornbook Magazine		X		
Human Factors	X			
Human Potential		X		
Human Relations (London)				Foreign
Industrial Arts and Vocational Education		X		
Industrial and Labor Relations Review			X	
Industrial Relations			X	
Industrial Training International		X		
Instructor		X		
Instrumentalist		X		
Integrated Education		X		
International Development Review		X		
International Journal of American Linguistics (IJAL)			X	
International Journal of Group Psychotherapy	X			
International Journal of Psychoanalysis (London)				Foreign
International Journal of Social Psychiatry (Britain)				Foreign

7.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L	NO EMPIR'L ON ED.	OTHER
IRAL- International Review of Applied Linguistics in Language Teaching				Foreign
International Review of Education (Holland)				Foreign
IRCE Bulletin (Information Retrieval Center on the Disadvantaged)		X		Newsletter
Journal of Abnormal Psychology	X			
Journal of the Acoustical Society of America	X			
Journal of Analytical Psychology (London)				Foreign
Journal of American History			X	
Journal of American Psychoanalytic Association	X			
Journal of Business Education				Articles too short (3 paragraphs)
Journal of Chemical Education	X			
Journal of Child Psychiatry	X			
Journal of Clinical Psychology			X	
Journal of College Placement	X			
Journal of Communication			X	
Journal of Comparative and Physiological Psychology (JCPP)	X			

8.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Journal of Conflict Resolution			X	
Journal of Cooperative Extension	X			
Journal of Education (Boston U.)		X		
Journal of Education for Librarianship		X		
Journal of Educational Administration (Australia)				Foreign
Journal of Educational Data Processing		X		
Journal of Experimental Psychology	X			
Journal of Experimental Social Psychology	X			
Journal of General Psychology	X			
Journal of Health and Social Behavior	X			
Journal of Higher Education		X		
Journal of the History of Ideas		X		
Journal of Human Relations		X		
Journal of Intergroup Relations	X			
Journal of Law and Economics			X	
Journal of Linguistics (London)				Foreign
Journal of Marriage and the Family	X			

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIRL.	NO EMPIR'L ON ED.	OTHER
Journal of Personality	X			
Journal of Politics			X	
Journal of Psychology	X			
Journal of Rehabilitation		X		
Journal of Research and Development in Education		X		
Journal of Secondary Education		X		
Journal of the SMPTE (Society of Motion Picture and Television Engineers)			X	
Journal of Social Psychology	X			
Journal on State School Systems Development		X		
Journal of Verbal Learning and Verbal Behavior	X			
Junior College Journal		X		
Kappa Delta Pi Record		X		
Kyklos				Foreign
Land Economics	X			
Language			X	
Language and Speech (Britain)				Foreign
Laryngoscope	X			
Library Journal			X	
Library Quarterly			X	
Library Resources and Technical Services			X	

10.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L	NO EMPIR'L ON ED.	OTHER
Lingua (Holland)				Foreign
Looking Ahead			X	
Main Currents in Modern Thought		X		
Management Science			X	
Media Methods		X		
Mental Retardation		X		
Midwestern Political Science Review			X	
Modern Language Quarterly				Discontinued
Monthly Labor Review (U.S. Bureau of Labor Statistics)			X	
Multivariate Behavioral Research	X			
Music Educator's Journal (MENC)		X		
Music Journal		X		
NAEB Journal (Educational Broadcasting Review)		X		
National Elementary Principal		X		
Nation's Schools		X		
NEA Journal		X		
NEA Reporter		X		
NEA Research Bulletin		X		
NEA Research Reports		X		It's "social bookkeeping," not research.

11.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
NSPI Journal (National Society for Programmed Instruction)			X	
Occupational Outlook Quarterly		X		
Orbis	X			
Organizational Behavior and Human Performance	X			
Pacific Affairs	X			
Pacific Sociological Review			X	
Perceptual and Motor Skills	X			
Personnel			X	
Personnel Journal		X		
Personnel Practice Bulletin (Australia)				Foreign
Personnel Psychology	X			
Phi Delta Kappan		X		
Physics Teacher			X	
PMLA - Publication of the Modern Language Association			X	
Political Science Quarterly			X	
Psychiatric Quarterly	X			
Psychoanalytic Quarterly	X			
Psychoanalytic Review	X			
Psychological Bulletin	X			
Psychological Reports	X			

Name	Reason for Elimination NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Psychological Review	X			Reviews
Psychology Today			X	
Psychometrika	X			
Psychonomic Science	X			
Public Administration Review	X	X		
Public Management		X		
Public Personnel Review	X	X		
Quarterly Journal of Economics	X			
Quarterly Journal of Experimental Psychology (England)				Foreign
Reading Improvement		X		
Reading News Report		X		
Reading Research Quarterly				Reviews
Religious Education		X		
Review of Educational Research				Review
Review of Politics	X			
Saturday Review		X		
School Administrator		X		
School Arts		X		
School Library Journal		X		
School Management		X		
School Musician		X		
School Science and Mathematics			X	

13.

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Science			X	
Science and Children		X		
Science and Society		X		
Science Teacher		X		
Social Education		X		
Social Psychiatry		X		
Social Research			X	
The Social Studies	X			
Sociological Quarterly			X	
Sociometry			X	
Southern Quarterly			X	
Southwestern Journal of Anthropology	X			
Soviet Education				Foreign
Speech Teacher			X	
Technical Education and Industrial Training (London)				Foreign
Technical Education News		X		
TESOL Quarterly (Teaching of English to Speakers of Other Languages)		X		
Theory Into Practice		X		
Think			X	
Training in Business and Industry		X		
Training and Development Journal		X		

Name	Reason for Elimination			
	NO EDUCA.	NO EMPIR'L.	NO EMPIR'L ON ED.	OTHER
Urban Affairs Quarterly	X			
Urban Education				Circula. too small (600)
Urban Review		X		
Volta Review		X		
Western Political Quarterly			X	
World Politics	X			
Yale Review		X		
Young Children		X		

APPENDIX II-3

Journals Scanned and Yielding Articles N = 113

	Number of Articles Yielded Altogether	After Duplicates Eliminated	In Sample to be Rated
Administrative Science Quarterly	6	6	2
Adult Education	4	4	1
American Behavioral Scientist	3	3	2
American Educational Research Journal	31	21	7
American Journal of Mental Deficiency	7	7	2
American Journal of Orthopsychiatry	3	3	2
American Journal of Psychiatry	2	2	1
American Journal of Sociology	3	2	0
American Political Science Review	1	1	0
American Sociological Review	3	2	1
American Sociologist	3	2	1
American Statistical Association Journal	1	1	1
Arithmetic Teacher	15	15	7
AV Communication Review	11	10	2
California Journal of Educational Research	22	18	6
Child Development	5	4	2
College and University	15	13	5
College Board Review	5	4	3
College Composition and Communication	2	2	0

	Number of Articles Yielded Altogether	After Duplicates Eliminated	In Sample to be Rated
College English	1	1	0
College Student Survey	2	2	1
Contemporary Education	3	2	0
Cornell Journal of Social Relations	1	1	0
Counselor Education and Supervision	12	9	3
Daedalus	1	1	0
Education	4	3	0
Educational and Psychological Measurement	12	11	3
Educational Record	3	3	0
Elementary School Guidance	1	1	0
Elementary School Journal	21	17	7
Exceptional Children	9	8	2
Foreign Language Annals	2	1	1
Fortune	1	1	1
Harvard Educational Review	4	4	0
Human Organization	1	1	1
Improving College and University Training	4	4	1
International Journal for the Education of the Blind	2	1	0
International Social Science Journal	1	1	0
Journal of Applied Behavior Analysis	6	6	2
Journal of Applied Behavioral Science	5	4	0

3

	Number of Articles Yielded Altogether	After Duplicates Eliminated	In Sample to be Rated
Journal of Applied Psychology	5	5	3
Journal of College Student Personnel	46	39	14
Journal of Consulting and Clinical Psychology	1	1	0
Journal of Counseling Psychology	45	37	14
Journal of Creative Behavior	2	2	0
Journal of Educational Measurement	25	21	5
Journal of Educational Psychology	39	33	14
Journal of Educational Research	24	20	6
Journal of Experimental Child Psychology	7	7	2
Journal of Experimental Education	48	45	18
Journal of General Education	1	1	0
Journal of Genetic Psychology	1	1	1
Journal of Human Resources	10	10	3
Journal of Learning Disabilities	1	1	0
Journal of the National Association of Women Deans and Counselors	3	2	1
Journal of Negro Education	8	7	3
Journal of Personality and Social Psychology	3	3	2
Journal of Political Economy	2	2	1
Journal of Projective Techniques and Personality Assessment	1	1	0
Journal of Reading	19	18	5
Journal of the Reading Specialist	10	7	2
Journal of Research in Music Education	12	12	3

	Number of Articles Yielded Altogether	After Duplicates Eliminated	In Sample to be Rated
Journal of Research in Science Teaching	30	28	8
Journal of School Psychology	13	12	7
Journal of Social Issues	5	5	2
Journal of Social Psychology	5	4	2
Journal of Special Education	11	10	3
Journal of Speech and Hearing Disorders	4	4	2
Journal of Speech and Hearing Research	2	2	1
Journal of Teacher Education	11	9	5
Language Learning	1	1	1
Liberal Education	2	2	0
Mathematics Teacher	3	3	1
Merrill Palmer Quarterly	5	5	3
Midwest Journal of Political Science	1	1	0
Modern Language Journal	9	8	4
NASPA	1	1	0
NASSP Bulletin (National Association of Secondary School Principals)	7	6	3
National Business Education Quarterly	3	3	1
National Catholic Education Association Bulletin	3	3	2
Peabody Journal of Education	5	5	3
Personnel and Guidance Journal	71	61	22
Phylon	2	2	0

5

	Number of Articles Yielded Altogether	After Duplicates Eliminated	In Sample to be Rated
Psychiatry	1	1	1
Psychology	2	2	2
Psychology in the Schools	28	22	8
Public Interest	2	2	2
Public Opinion Quarterly	2	2	0
Quarterly Journal of Speech	2	2	1
Reading Teacher	26	26	10
Research in the Teaching of English	7	6	1
Review of Economics and Statistics	2	2	0
Rural Sociology	1	0	0
School and Society	3	3	0
School Counselor	16	15	5
School Review	10	8	1
Science Education	18	17	8
Scientific American	1	1	0
Social Forces	1	1	0
Social Problems	1	1	0
Sociological Analysis	1	1	0
Sociological Inquiry	2	2	0
Sociology and Social Research	3	3	0
Sociology of Education	20	16	5
Southern Economic Journal	1	1	0

6

	Number of Articles Yielded Altogether	After Duplicates Eliminated	In Sample to be Rated
Southern Journal of Educational Research	11	10	3
Speech Monographs	12	11	4
Studies in Art Education	10	10	6
Teachers College Journal (Ind. U.)	3	3	1
Teachers College Record	5	5	2
Training School Bulletin	5	5	2
Trans-Action	2	2	1
Vocational Guidance Quarterly	17	16	6
Total	960	848	303

APPENDIX II-4

COVER LETTER TO AUTHORS ACCOMPANYING QUESTIONNAIRE

Columbia University in the City of New York | *New York, N.Y. 10025*

BUREAU OF APPLIED SOCIAL RESEARCH

605 West 115th Street

(Date)

(Name of Author) *
Title
Address

Dear (name of author):

The U.S. Office of Education has asked us to gather some information about colleagues who published a research article on education in 1967-68. Your article which was selected by our sampling procedures is:

(name of research article)

The Office of Education is concerned about the extent to which current output contributes to basic knowledge or to the improvement of education. In addition, it is eager to improve research on education, and therefore your experiences are very important to them.

Many people are trained to do research but relatively few publish. The fact that you have recently published your research makes us interested in your background and how you decided to do the research you did. We plan to have a panel of readers describe the procedural and substantive features of your article. This information will then be related to your background and experiences.

We are sure that you, like all of us, have been inundated by questionnaires. But since the current survey might serve as a basis for changes in the policies and procedures of research funding, we urge you to cooperate. The questionnaire can be mailed without postage by stapling or taping the open edge. Thank you very much.

Sincerely yours,

Sam D. Sieber

Sam D. Sieber
Program Director

Paul F. Lazarsfeld
Faculty Consultant

Caroline Hodges Persell

(Mrs.) Caroline Hodges Persell
Research Manager

*(parenthetical items were individually typed in matching type)

QUESTIONNAIRE FOR AUTHORS OF RESEARCH ON EDUCATIONAL TOPICS

Supported by
The U.S. Office of Education

Bureau of Applied Social Research
Columbia University
605 West 115 Street
New York, N.Y. 10025

Budget Bureau No. 51-S-69007
Approval Expires 4/30/70 **1-2/02**

THIS QUESTIONNAIRE IS AN ESSENTIAL PART OF A LARGER STUDY OF THE TRAINING, CAREER LINES, TOPIC SELECTION, AND RESEARCH OUTPUT OF PEOPLE WHO HAVE DONE EMPIRICAL RESEARCH ON SOME ASPECT OF EDUCATION. Feel free to add comments or explanations at any point. If you have difficulty remembering any of the information requested, your best estimate will still be very helpful.

The information requested in this questionnaire is regarded as confidential and will be used for statistical purposes only. It will NOT be released in any way that will allow it to be identified with you.

The first part of the questionnaire deals with your paper entitled

I. PAPER RELATED QUESTIONS

3-6/

A. Topic of the Paper

Would you please think about the origin of the topic of your paper mentioned above. By topic we mean the specific idea or problem around which the paper is written. (We do not mean the general theoretical orientation which you may bring to bear on most of the problems you consider, or the empirical project on which this paper is based. As an illustration, your general theoretical orientation might be learning theory and your empirical project might be learning in certain types of classroom situations, with the specific topic of your paper being a comparison of learning in a stressful teaching situation versus learning in a more supportive classroom environment.)

1.1 How did the topic of the paper mentioned above come to your attention?

(Please do two things:
(1) check all that apply
and
(2) circle the number of the single most important source)

01 ☐ colleagues where you work
02 ☐ other colleagues
03 ☐ fellow graduate students
04 ☐ your professor(s) or former professor(s)
05 ☐ your student(s)
06 ☐ client or funding agent
07 ☐ administrative superior
08 ☐ emphasis of a research program in an organization where you work
09 ☐ reading in theoretical literature
10 ☐ reading in empirical literature
11 ☐ questions or ideas from your own previous research
12 ☐ your own reflection, i.e., not based on previous research or others' ideas
13 ☐ Educational Resources Information Centers (ERIC) or publications
14 ☐ your own field observation of a concrete problem in education
15 ☐ teachers or other practitioners in school systems
16 ☐ unpublished research
17 ☐ general interest of educators in the problem
18 ☐ emphasis of the U.S. Office of Education on the problem
19 ☐ non-professional friend or acquaintance
20 ☐ other (please specify)____________________

____

7-8/

1.2 What was it about this topic which attracted you?

(Please do two things:
(1) check all that apply
and
(2) circle the number of the single most important feature)

01 ☐ important for solving a concrete problem in education
02 ☐ good for testing a particular research method
03 ☐ feasible or manageable
04 ☐ of interest to a funding agent
05 ☐ important for developing or testing an educational theory
06 ☐ important for developing or testing a behavioral science theory
07 ☐ acceptable to your advisor or administrative superior
08 ☐ personally interesting
09 ☐ opportunity for exploring a new area
10 ☐ other (please specify) ____________________
____

9-10/

B. The Paper in General

1.3 When did you finish writing this paper? ____________________ (Please indicate the approximate month and year) 11-14/

1.4 Was the paper part of your master's thesis, doctoral dissertation, or a student paper?

1 ☐ master's 2 ☐ doctorate 3 ☐ student paper 4 ☐ none of these 15/

C. Empirical Project on which the Paper is Based

1.5 Was any of the research in this project done in or through a research organization or bureau?

1 ☐ yes 2 ☐ no 16/

IF YES, please give the name and location of that research bureau or organization:

____________________ (name)
____________________ (location)

1.6 What was the total cost of the research project (including overhead, salaries paid to researchers and clerical staff, data processing, field expenses, travel, equipment, etc.)?

$__________ (approximately) 17-22/

1.7 Did the project receive financial support for research expenses from university funds, a foundation, a governmental agency, etc.?

1 ☐ yes 2 ☐ no (IF NO, please skip to Q. 1.8) 23/

IF YES, from what source? ____________________

IF FROM THE FEDERAL GOVERNMENT, from what agency? 24-25/

(Please do two things:
(1) check all that apply
and
(2) circle the number of the single most important source)

01 ☐ Children's Bureau
02 ☐ Department of Defense
03 ☐ Department of Labor
04 ☐ National Aeronautics and Space Agency
05 ☐ National Institutes of Health
06 ☐ National Institute of Child Health and Human Development
07 ☐ National Institute of Mental Health
08 ☐ National Science Foundation
09 ☐ Office of Education
10 ☐ Office of Economic Opportunity
11 ☐ Office of Naval Research
12 ☐ Public Health Service
13 ☐ Other (please specify) ____________________

____

1.8 Was the research on which this paper is based part of a larger research project?

1☐ yes 2☐ no 26/

Your Position while Doing this Research

1.9 Did you direct this research project (i.e., have primary responsibility for its design, execution, and analysis)?

1☐ yes 2☐ no 27/

1.10 At the time of this research,

A. What was your main institutional affiliation?

1☐ University
2☐ College
(a) ____________________ (name) ____________________ (city and state) 28/

(b) What school or department were you in? 29/

School or Department of Education
1 ☐ undergraduate
2 ☐ graduate
3 ☐ joint undergraduate/graduate

Liberal arts department (please specify discipline) ____________________
4 ☐ undergraduate
5 ☐ graduate
6 ☐ joint undergraduate/graduate

Other school or department (please specify) ____________________
7 ☐ undergraduate
8 ☐ graduate
9 ☐ joint undergraduate/graduate

3☐ School System ____________________ (name) ____________________ (city and state)

4☐ State Department of Education ____________________ (state)

5☐ Private Agency ____________________ (name) ____________________ (city and state)

6☐ Other (please specify) ____________________ (name) ____________________ (city and state)

B. What was your principal position at this institution? (e.g. professor of educational psychology) 30-

____
(title of position)

1.11 At that time, were you a staff member of a research organization or bureau?

1 ☐ yes 2 ☐ no 32.

IF YES, please give the name and location of that research bureau or organization if it is different from that in question 1.5:

________________________ ________________________
(name) (location)

1.12 At the time of this research, how did you divide your professional time? (Please give your best estimate of the percent of time you spent on each activity.)

% of time	Activity	
_____	Teaching	33-
_____	Research	35-
_____	Service (e.g., consultation, developing or implementing new curriculum or action programs, school surveys, test administration or scoring, workshops.)	37-
_____	Other (please specify, e.g., administration, studying.) __________	39-
100%	________________________	41/

II. CAREER ACTIVITIES AND OPINIONS

A. Your Present Position

2.1 At the present, how do you divide your professional time? (Please give your best estimate of the percent of time you spend on each activity.)

% of time	Activity	
_____	Teaching	42-
_____	Research	44-
_____	Service (e.g., consultation, developing or implementing new curriculum or action programs, school surveys, test administration or scoring, workshops.)	46-
_____	Other (please specify, e.g., administration, studying.) __________	48-
100%	________________________	50/

2.2 At the present,

A. What is your main institutional affiliation? 51/

1 ☐ University
2 ☐ College
(a) ______________________ (name) ______________________ (city and state)

(b) What school or department are you in? 52/

School or Department of Education
1 ☐ undergraduate
2 ☐ graduate
3 ☐ joint undergraduate/graduate

Liberal arts department (please specify discipline) ______________________
4 ☐ undergraduate
5 ☐ graduate
6 ☐ joint undergraduate/graduate

Other school or department (please specify) ______________________
7 ☐ undergraduate
8 ☐ graduate
9 ☐ joint undergraduate/graduate

3 ☐ School System ______________________ (name) ______________________ (city and state)

4 ☐ State Department of Education ______________________ (state)

5 ☐ Private Agency ______________________ (name) ______________________ (city and state)

6 ☐ Other (please specify) ______________________ (name) ______________________ (city and state)

B. What is your principal position at this institution? (e.g., professor of educational psychology)

____ 53-54/
(title of position)

2.3 Please rank three of the following activities in terms of your personal preference (write 1 next to the job you would like to do most, 2 next to your second choice, and 3 next to your third choice).

55-56/^{1}
57-58/^{2}
59-60/^{3}

01____teaching elementary or secondary school students
02____teaching undergraduate students
03____teaching graduate students
04____educational administration
05____college or university administration
06____empirical research
07____research related activities (e.g., research administration)
08____consulting
09____service (e.g., developing or implementing new curriculum or action programs, school surveys, test administration or scoring, workshops)
10____in-service training of teachers
11____translating research into educational practice
12____other (please specify) ______________________
____

2.4 What is your major field or specialty? (Please check one only): 61-62/

01☐ curriculum
02☐ educational administration
03☐ guidance and counseling
04☐ higher education
05☐ subject matter area (e.g., art, business, language, math, music, science, etc.)
06☐ teacher training
07 ☐ testing and measurements
08 ☐ educational psychology
09 ☐ other psychology
10 ☐ sociology
11 ☐ other (please specify what) ______________________

2.5 Are you currently a staff member of a research organization or bureau? 63/

1 ☐ yes 2 ☐ no

IF YES, please give the name and location of that research bureau or organization.

______________________ (name) ______________________ (location)

2.6 Are you currently engaged in research?

1 ☐ yes 2 ☐ no (IF NO, please skip to Q. 2.7) 64/

IF YES, is the research related to education? 65/

1 ☐ yes 2 ☐ no

B. Your Opinions

2.7 Some researchers interested in education seek mainly to achieve recognition from behavioral scientists outside the field of education, while others are primarily concerned with being recognized by researchers within education or by school practitioners. Please check the group whose judgment is most important to you personally. (Please check one only)

1 ☐ researchers within education 2 ☐ researchers outside of education 3 ☐ school practitioners 66/

2.8 Thinking about the present state of research in education, do you think it should mainly: (Please check one only)

1 ☐ contribute directly to the solution of concrete problems in education 67/
2 ☐ contribute to the formulation or development of theory in a discipline which may or may not help solve problems in education. 1-2

3-6

2.9 Of course, many factors motivate individuals to do research. Regarding your usual research motives, how would you distribute 10 points among the following (so the total equals 10)?

7/____curiosity about substantive questions in education
8/____curiosity about substantive questions in a behavioral science field
9/____desire to contribute to the solution of concrete educational problems
10/____interest in research procedures
11/____requirements of present employment
12/____degree requirements
13/____desire for career advancement
14/____desire for prestige
15/____desire to earn more money
16/____desire to avoid the "firing line" of school teaching or administration
17/____other (please specify) ______________________

______________________ 18-

C. Your Activities

2.10 Have you ever taught? 1 ☐ yes 2 ☐ no 20/

IF YES,

a) How many years have you taught in elementary or secondary school?_________ 21-22/

b) How many years have you taught college undergraduates?_________ 23-24/

c) How many years have you taught graduate students?_________ 25-26/

d) How many years have you taught elsewhere; and in what kind of setting? _________ (years) _________ (type of setting) 27-28/

2.11 Other than while you worked on your master's or doctoral thesis, has there ever been a period of six months or more during which research was your primary activity?

1 ☐ yes 2 ☐ no (IF NO, please skip to Q. 2.12) 29/

IF YES, about how much time have you spent altogether when research was your primary activity?______years 30-31/

2.12 Career history:

For each of the major positions or ranks you have held since obtaining your highest degree, please indicate:

(a) the number of years you held the position or rank, and
(b) your activities in each position, ranking them according to the usual amount of time you devoted to each activity.

(Please consider it a new position whenever you changed title or rank within the same institution, or when you moved to a new institution.)

POSITIONS HELD OVER TIME	(a) NUMBER OF YEARS IN EACH POSITION	(b) RANKING OF ACTIVITIES (Rank 1 to 4 for each activity, with 1 as the activity you spent the most time doing, etc. Use 0 if you spent no time on an activity.)				
		(1) TEACHING	(2) RESEARCH	(3) SERVICE *	(4) OTHER **	
EXAMPLE:	2	1	3	2	0	
first						32/
second						33/
third						34/
fourth						35/
fifth						36/
sixth						
seventh						
eighth						

Please use additional space if necessary.

---

* Examples of service are consultation, developing or implementing new curriculum or action programs, school surveys, test administration or scoring, workshops.

** This category includes all your other professional activities, such as administration, studying, staff or committee meetings, etc.

2.13 Have you ever applied for a research grant from the U.S.O.E., N.I.M.H., Ford, Rockefeller, or Carnegie foundations, or any other such organization that supports research?

1 ☐ yes 2 ☐ no (IF NO, please skip to Q. 2.14) 37/

IF YES, have you ever received a research grant from such an organization?

1 ☐ yes 2 ☐ no (IF NO, please skip to Q. 2.14) 38

IF YES, please name the source; ____________________

2.14 Please list the major national professional associations to which you belong, using identifying words in full:

____

____

____________________ 39-

____

____

____

2.15 How many empirical research studies (articles, monographs, or books) have you published in addition to the one referred to on page one, and what was the date of your first publication? 41/

______number of articles ______number of monographs ______number of books 42/

43/

Date of first publication: ____________________ 44-

(month and year)

III. EDUCATIONAL BACKGROUND

A. General Background

3.1 Please list the colleges and universities you have attended:

INSTITUTION	STATE	NAME OF DEGREE	YEAR OF DEGREE	MAJOR FIELD	Was degree in Education? YES (1)	NO (2)

48/

49/

50-

3.2 As an undergraduate, did you receive any academic honors? (Please check all that apply)

HONOR	YES
Phi Beta Kappa	☐
Cum Laude	☐
Magna Cum Laude	☐
Summa Cum Laude	☐
Dean's List	☐
Other honors (please specify) ____________________	☐

52/

____

3.3 If you know them, what were your Graduate Record Exam (GRE) scores?

	VERBAL	MATHEMATICS	YEAR YOU TOOK THEM	
Absolute score	________	________	________	53-55/
Percentile	________	________	________	56-58/

8 ☐ Don't know them 9 ☐ Never took GRE's 59/

B. Experiences in Graduate School

(Please check here if you did not attend graduate school, and skip to Question 4.1.)

1 ☐ I did not attend graduate school. 60/

3.4 In your first year of graduate school, did you have a fellowship (other than the GI Bill) which required no work from you?

1 ☐ yes 2 ☐ no 61/

3.5 While in graduate school, did you ever work:

	(1) YES	(2) NO	
(a) as a teaching assistant	☐	☐	62/
(b) as a research assistant	☐	☐	63/
(c) in a non-professional position	☐	☐	64/

3.6 If you were a research assistant (IF NOT, please skip to Q. 3.7), what were your major activities?

(Please do two things:
(1) check all that apply,
and
(2) circle the number of the single activity you did most.)

1 ☐ Clerical work—typing, collating, filing, administrative details, etc.

2 ☐ Coding

3 ☐ Computing statistics

4 ☐ Computer programming, or writing assemblies for computer

5 ☐ Bibliographic and library work

6 ☐ Abstracting and synthesizing pertinent research

7 ☐ Field work (e.g., observation, interviewing, running tests on subjects, etc.)

8 ☐ Analytical work (e.g., helping in the study's conception, design, analysis, etc.)

9 ☐ Other (please specify) ________________

____

65/

3.7 As a graduate student did you ever work in a research center or bureau within the university?

1 ☐ yes 2 ☐ no 66/

IF YES, please give the name and location of that center or bureau:

________________ (name)

________________ (location)

C. Course Work In Graduate School

3.8 Approximately how many graduate semesters, trimesters, or quarters did you take in each of the following areas? (Please check the appropriate boxes indicating the number of semesters, trimesters or quarters you took in groups A through G.)

	NUMBER OF GRADUATE SEMESTERS, TRIMESTERS, OR QUARTERS TAKEN					
	None (0)	One (1)	Two (2)	Three (3)	More than three (4)	
A. COURSES IN PROFESSIONAL EDUCATION e.g., educational administration, higher education, guidance and counseling, curriculum and teaching in elementary or secondary school, subject matter courses (art, business, language, math, music, science, speech education, etc.)	☐	☐	☐	☐	☐	67/
B. COURSES IN PSYCHOLOGY e.g., educational psychology, child development, experimental psychology, learning, etc.	☐	☐	☐	☐	☐	68/
C. COURSES IN SOCIAL SCIENCE e.g., anthropology, economics, government or political science, sociology	☐	☐	☐	☐	☐	69/
D. COURSES IN HUMANITIES e.g., history, literature, philosophy, foreign language, etc.	☐	☐	☐	☐	☐	70/
E. COURSES IN RESEARCH METHODS e.g., experimental design, field observation, social survey design and analysis, testing and measurements, etc.	☐	☐	☐	☐	☐	71/
F. COURSES IN RESEARCH AIDS						
computer programming	☐	☐	☐	☐	☐	72/
statistics	☐	☐	☐	☐	☐	73/
other (please specify) ______________	☐	☐	☐	☐	☐	74/
G. OTHER TYPES OF COURSES (Please specify)						
______________	☐	☐	☐	☐	☐	75/
______________	☐	☐	☐	☐	☐	76/

1-2/04

3-6/

3.9 As a graduate student, were you ever in a special program for training people to do research on education, other than the regular graduate degree program?

1 ☐ yes 2 ☐ no (IF NO, please skip to Q. 4.1) 7/

a) IF YES, to the best of your knowledge, was this a U.S. government-sponsored program?

1 ☐ yes 2 ☐ no 8/

b) IF YES to (a), did you personally receive financial support from this program?

1 ☐ yes 2 ☐ no 9/

IV. GENERAL BACKGROUND

4.1 Birthdate: ________ (month) ________ (year) 10-13/

4.2 Sex: 1 ☐ male 2 ☐ female 14/

4.3 Race: 1 ☐ white 2 ☐ Negro 3 ☐ Other (please specify) ____________________ 15/

4.4 Religion (optional): 16/

1 ☐ Catholic
4 ☐ other (please specify) ____________________
____
2 ☐ Jewish
3 ☐ Protestant (please specify denomination ____________________)
5 ☐ none

4.5 What was your marital status at the time you wrote the paper mentioned on page one? 17/

1 ☐ single
2 ☐ divorced
3 ☐ married
4 ☐ widowed
5 ☐ separated

4.6 How many children did you have at that time? ________ 18-19/

4.7 Please indicate by check marks the highest educational attainment of each parent. 20/ 21/

		(20) Father		(21) Mother
Elementary school	1	________	1	________
Some high school	2	________	2	________
High school graduate	3	________	3	________
Some college	4	________	4	________
College graduate	5	________	5	________
Some graduate study	6	________	6	________
Master's Degree	7	________	7	________
Doctor's Degree	8	________	8	________

4.8 In which category below was your total income before taxes in 1968? 22/

1 ☐ Under $5,000
2 ☐ $5,000–7,499
3 ☐ $7,500–9,999
4 ☐ $10,000–14,999
5 ☐ $15,000–19,999
6 ☐ $20,000–24,999
7 ☐ $25,000–29,999
8 ☐ Over $30,000

4.9 Approximately what per cent of your income came from: 23-24/

teaching salary ________% 25-26/
royalties on books ________%
research salary ________% 27-28/
other (please specify) ______%
consulting fees ________% 29-30/
________________________________ 31-32/
consulting fees ________% 33-34/

35/

36/

THANK YOU FOR YOUR COOPERATION IN COMPLETING THIS QUESTIONNAIRE. PLEASE STAPLE OR TAPE THE OPENING EDGE AND MAIL. WE WILL PAY THE POSTAGE.

THANK YOU VERY MUCH.

FIRST CLASS
Permit No. 55988
New York, N.Y.

BUSINESS REPLY MAIL

No Postage Stamp Necessary if Mailed in the United States

Postage will be Paid by:

Bureau of Applied Social Research
Columbia University
605 West 115th Street
New York , N. Y. **10025**

Attention: DEW

STAPLE HERE

APPENDIX II-6
ONE PAGE QUESTIONNAIRE
FOR AUTHORS OF RESEARCH ON EDUCATIONAL TOPICS

1-2/08 In reference to your paper entitled:

3-6/

1. Was there a special reason why you did not fill out and return the "Questionnaire for Authors of Research on Educational Topics"?

(Please do two things:
(1) check all that apply
and
(2) circle the number of the single most important reason.)

1___I do not consider the paper mentioned a research paper.
7/ 2___I do not think my research was on an "educational topic".
3___I have not done any research since, nor do I plan to do any in the future.
4___I am only peripherally interested in education.
5___The questionnaire was too long.
8/ 6___I do not believe in survey research.
7___Other (please specify what_____________________________)

2. Currently what is your major field or specialty? (Please check one only):

9-10/
01___curriculum
02___educational administration
03___guidance and counseling
04___higher education
05___subject matter area (e.g., art, business, language, math, music, science, etc.)
06___teacher training
07___testing and measurements
08___educational psychology
09___other psychology
10___sociology
11___other (please specify what) ________________

3. What is your main institutional affiliation?

11/ 1___University
2___College ____________________ ____________________
(name) (city and state)

3___School System________________(name)______________(city and state)

4___State Department of Education__________________________(state)

5___Private Agency________________(name)______________(city and state)

6___Other (please specify)____________(name)____________(city and state)

(OVER PLEASE)

APPENDIX II-7

CATEGORIZATION OF JOURNALS BY TYPE

Education

Adult Education
American Educational Research Journal
Arithmetic Teacher
California Journal of Educational Research
College and University
College Board Review
College Composition and Communication
College English
College Student Survey
Contemporary Education
Counselor Education and Supervision (ACES)
Education
Educational and Psychological Measurement
Educational Record
Elementary School Guidance
Elementary School Journal
Exceptional Children
Foreign Language Annals
Harvard Educational Review
Improving College and University Training
Journal of College Student Personnel
Journal of Educational Measurement
Journal of Educational Research
Journal of Experimental Education
Journal of General Education
Journal of Learning Disabilities
Journal of the National Association of Women Deans and Counselors
Journal of Negro Education
Journal of Reading
Journal of the Reading Specialist
Journal of Research in Music Education
Journal of Research in Science Teaching
Journal of Special Education
Journal of Teacher Education
Language Learning
Liberal Education
Mathematics Teacher
Merrill Palmer Quarterly

Modern Language Journal
NASPS
NASSP Bulletin
National Business Education Quarterly
National Catholic Education Association Bulletin
Peabody Journal of Education
Personnel and Guidance Journal
Reading Teacher
Research in the Teaching of English
School and Society
School Counselor
School Review
Science Education
Southern Journal of Educational Research
Studies in Art Education
Teachers College Journal (Indiana University)
Teachers' College Record
Training School Bulletin
Vocational Guidance Quarterly

Behavioral Science Discipline

Administrative Science Quarterly
American Behavioral Scientist
American Journal of Orthopsychiatry
American Journal of Psychiatry
American Journal of Psychology
American Journal of Sociology
American Political Science Review
American Sociological Review
American Sociologist
American Statistical Association Journal
Child Development
Cornell Journal of Social Relations
Human Organization
International Social Science Journal
Journal of Applied Behavior Analysis
Journal of Applied Behavioral Science
Journal of Applied Psychology
Journal of Consulting and Clinical Psychology
Journal of Counseling Psychology
Journal of Educational Psychology
Journal of Experimental Child Psychology
Journal of Genetic Psychology
Journal of Human Resources
Journal of Personality and Social Psychology
Journal of Political Economy
Journal of Projective Techniques and Personality Assessment

Journal of School Psychology
Journal of Social Issues
Journal of Social Psychology
Midwest Journal of Political Science
Psychiatry
Psychology
Psychology in the Schools
Public Opinion Quarterly
Review of Economics and Statistics
Rural Sociology
Social Forces
Social Problems
Sociological Analysis
Sociological Inquiry
Sociology and Social Research
Sociology of Education
Southern Economic Journal
Transaction

General

Daedalus
Fortune
Phylon
Public Interest
Scientific American

Other

American Journal of Mental Deficiency
Audio-Visual Communication Review
International Journal for the Education of the Blind
Journal of Creative Behavior
Journal of Speech and Hearing Disorders
Journal of Speech and Hearing Research
Speech Monographs

APPENDIX IV-1

DETAILED CRITERIA RATING FORM

Developed by Caroline Hodges Persell

(Revised Version No. 2, July 1966)

Rater's Name: ______________________________

Article No.: ______________

Date: ____________________

Section A: STATEMENT OF AND JUSTIFICATION FOR ORIENTING IDEAS

A 1 Delineation of Initial Concepts (or major dimensions of study):

A 1.1 Conceptual terms are clear, including technical ones.

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

A 1.2 Presentation is in orderly, logical sequence.

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

A 2 **Specification of Objective** (specific problems or hypotheses to be investigated):

A 2.1 Objectives are stated explicitly.

1-No 2-Questionable 5-Yes

A 2.2 Rationale for pursuing the objectives is indicated (DNA if objectives not stated).

0-DNA 1-No 5-Yes

A 2.3 The rationale for pursuing the objectives is justified (DNA if objectives not stated).

0-DNA

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

A 3 **Relationship to existing knowledge**:

A 3.1 Relationship to existing knowledge is indicated (e.g., extension, specification, filling in gaps, etc.).

1-No 3-Questionable 5-Yes

A 3.2 Relationship to existing knowledge is apt (DNA if relationship is not stated).

0-DNA

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

A 4 Significance of questions raised

A 4.1 Answers to questions raised would make a significant contribution to a theoretical problem.

1-Not true 3-Somewhat true 5-Very true

Section B: RESEARCH DESIGN AND EXECUTION

B 1 Description of design:

B 1.1 The nature of each "case" is described sufficiently for replication of the study.

1-No 3-Questionable 5-Yes

B 1.2 The measuring devices (indicators, indices, etc.) representing the major concepts:

B 1.2a Are described.

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

B 1.2b Are valid (i.e., neither exceed nor under-represent the concepts).

0-DNA (not described)

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

B 1.3 The population to which the sample or case may be generalized:

B 1.3a Is described or is clearly obvious.

0-DNA (if non-sample study) 1-No 5-Yes

B 1.3b Is an appropriate population in terms of the study's objectives (DNA if the population is not clear).

0-DNA 1-No 5-Yes

B 1.4 The size of the sample is stated (DNA if non-sample study).

0-DNA 1-No 5-Yes

B 1.5 Any sub-group comparisons which are built into the design, and the logical bases for these comparisons, are described (as in experiments, longitudinal or panel studies, stratified survey samples, or comparative case studies).

0-DNA (sub-groups not built in)

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

B 2 Execution of the study design:

B 2.1 In the collection of the data:

B 2.1a The initial sample was representative.

1-No 3-Questionable 5-Yes

B 2.1b Measures were taken to insure maximum response rate from the sample or, as in longitudinal studies, to keep differential loss of subjects to a minimum.

1-No 3-Questionable 5-Yes

B 2.1c The type of sample or selection procedure utilized was appropriate to the objectives of the study (e.g., cluster sampling, stratified sampling, snow-ball sampling, frequency distribution matching, etc.).

1-No

2-Less than adequate

3-Questionable

4-More than adequate

5-Yes

B 2.1d The following problems were dealt with where necessary.

1) Statistical regression effect (results when groups are selected on the basis of their extreme scores or positions).

0-DNA 1-Inadequate 3-Questionable 5-Adequate

2) Selection bias (differential selection of comparison groups).

0-DNA 1-Inadequate 3-Questionable 5-Adequate

3) Interaction effects of selection biases and the experimental variable.

0-DNA 1-Adequate 3-Questionable 5-Inadequate

4) Instrumentation effect (changes in question wording, interviewing behavior, etc.).

0-DNA 1-Inadequate 3-Questionable 5-Adequate

B 2.1e If the research design longitudinal, panel, or experimental?

1-No* 5-Yes**

* IF "No" -- skip to section B 2.1f.

** IF "Yes" -- Are the following possible sources of unreliability dealt with?

1) Maturation effect (changes due to passage of time).

1-Inadequate 3-Questionable 5-Adequate

2) Testing effect (experience of earlier test affects results of later test through cueing, etc.).

1-Inadequate 3-Questionable 5-Adequate

3) Sensitizing effect of pre-test or first interview (affects subject's responsiveness to subsequent stimuli or events).

1-Inadequate 3-Questionable 5-Adequate

4) Artificiality of setting or subject's knowledge that he is participating in an experiment.

1-Inadequate 3-Questionable 5-Adequate

5) Multiple treatment effect (caused whenever multiple treatments are applied to the same subjects because effects of prior treatments are not usually erasable).

1-Inadequate 3-Questionable 5-Adequate

6) Possibility of selection-maturation interaction, selection-testing interaction, or selection-instrumentation interaction (interactions between selection bias and any of the above which might be mistaken for effect of experimental variable).

1-Inadequate 3-Questionable 5-Adequate

B 2.1f Any other possible weaknesses in the design:

1) Are acknowledged.

0-DNA (no others) 1-No 3-Questionable 5-Yes

2) Are dealt with.

0-DNA (if not acknowledged)

1-Poorly

2-Less than adequately

3-Adequately

4-More than adequately

5-Outstandingly

B 2.2 Analytical strategy of the author:

B 2.2a Data is organized according to major concepts which have been introduced --

0-DNA

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

B 2.2b Principal variables which have been measured are handled in such a way as to --

1) Reveal distributions, relationships, effects, or comparisons bearing on the hypotheses or problems.

0-DNA

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

2) Discount uncontrolled factors which might be affecting the results (or spurious relationships resulting from a third variable which independently affects both variables in a given relationship).

0-DNA (if matching was done; if descriptive; if controls for spuriousness not necessary)

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

3) Specify relationships between two or more variables by showing the conditions under which the relationships are stronger or weaker, where needed.

0-DNA (if matching was done; if descriptive; if specifications not necessary)

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

B 2.3 Analytical techniques of the author.

B 2.3a Use of statistics:

1) Statistical techniques are appropriate to the purposes of the study (e.g., regression analysis, analysis of variance, factor analysis, percentage differences, etc.).

0-DNA

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

2) Tests of significance are used appropriately for the sample design and for the nature of the data. IF NOT USED: Decision not to use them is justified by the purpose or design of the study (e.g., they would not be justified when probability tests were used in conjunction with non-random samples and non-normal distributions).

0-DNA 1-No 3-Questionable 5-Yes

B 2.3b Qualitative analysis (e.g., informant interviewing, observation, case studies, documentary analysis, etc.):

1) Qualitative analysis is used where needed.

0-DNA (not needed) 1-No 3-Questionable 5-Yes

2) Qualitative analysis is carried out properly.

0-DNA (if not used)

1-Poorly

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

Section C: PRESENTATION OF DATA

C 1 Scope of presentation:

C 1.1 All of the data specified in the design as being relevant to the study's objectives is presented or summarized.

1-No 3-Questionable 5-Yes

C 2 Clarity of presentation:

C 2.1 Topics or points are organized in a logical, coherent sequence.

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

C 2.2 Technical terms are appropriate (unnecessary jargon is avoided, usage is consistent, and new terms are clearly defined).

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

Section D: APTNESS OF INTERPRETATIONS

D 1 Conclusions are appropriate to the data.

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

D 2 Data make a useful contribution to the theoretical or practical problems posed by the research.

1-Poor

2-Less than adequate

3-Adequate

4-More than adequate

5-Outstanding

D 3 Data which do not support the hypotheses, common expectations, or previous findings are acknowledged.

1-Never

2-Seldom

3-Sometimes

4-Usually

5-Always

D 4 Implications for future research are mentioned.

1-No 3-Questionable 5-Yes

If any of the above dimensions (items or sections) were so poorly handled as to invalidate most of the report, or so well handled that this report was truly outstanding, please indicate the number of the dimension below:

Poorly handled ____________________

Well handled ______________________

This rating took:

Less than 15 minutes _________

15 to 30 minutes _____________

31 to 45 minutes _____________

46 to 60 minutes _____________

More than 60 minutes _________

Signed:______________________________________

APPENDIX IV-2

LETTER TO JUDGES

Columbia University in the City of New York | *New York, N.Y. 10025*
BUREAU OF APPLIED SOCIAL RESEARCH
605 West 115th Street

April 18, 1969

(Name and address)

Dear (name of judge):

We are writing to ask if you would be willing to judge the quality of 11 journal articles primarily from your specialty of (name of specialty), using the enclosed rating form. The articles average five to ten pages in length. For doing the ratings, you will receive a $40 honorarium.

This rating is part of a study we have been asked by the U.S. Office of Education to do on the training, career lines, topic selection, and research quality of people studying some aspect of education. A questionnaire is being sent to 1,100 authors of empirical research articles on education. Four hundred and forty articles by these authors will be rated by 40 judges. We urge you to cooperate because the current survey may serve as a basis for changes in the policies and procedures of research funding. The final worth of the study depends upon having the articles rated by qualified judges such as yourself. Enclosed is an initial list of people we are asking to be judges. We anticipate some changes in this list.

Would you please use the enclosed postcard to let us know whether or not you will rate the 11 articles. If you agree, we will send you rating forms and Xeroxed copies of the articles. We will need the completed ratings within about four weeks after you receive the materials. We hope you will be able to work with us.

Sincerely yours,

Sam D. Sieber
Program Director

Paul F. Lazarsfeld
Faculty Consultant

(Mrs.) Caroline H. Persell
Research Manager

Enc. Reply postcard
Initial list of judges
Sample rating form

Columbia University in the City of New York | *New York, N.Y. 10025*

BUREAU OF APPLIED SOCIAL RESEARCH

605 West 115th Street

APPENDIX IV-3

INSTRUCTION LETTER TO JUDGES

(date)

(Name and address)

Dear (name):

We are delighted you have agreed to rate research papers for our study of educational researchers. Enclosed are xeroxed copies of eleven papers, plus Rating Forms and brief Specification Forms for each paper. Also enclosed is a self-addressed, stamped envelope in which you can return the Rating and Specification Forms to us.

After reading each article please rate it on the Rating Form and check the three questions on the Specification Form. Please be sure the four digit number appearing in the upper right hand corner of the article is added to the Rating and Specification Forms.

We hope you will be able to finish all of the ratings within three or four weeks, since we must adhere to a tight schedule.

Thank you very much again for your cooperation.

Sincerely yours,

Sam D. Sieber
Program Director

Paul F. Lazarsfeld
Faculty Consultant

(Mrs.) Caroline H. Persell
Research Manager

APPENDIX IV-4

DISTRIBUTION OF JUDGES' RATINGS BY JUDGE

		Rating			
Judge #	Theory Practice Methods*	Above Average	Average	Below Average	Total
1	T	5	4	0	9
	P	7	0	2	9
	M	1	2	5	8
2	T	1	4	5	10
	P	1	5	4	
	M	1	5	4	
3	T	2	3	4	9
	P	4	2	3	=
	M	4	2	3	
4	T	2	2	5	9
	P	3	1	5	
	M	3	1	5	
5	T	5	3	2	10
	P	4	4	2	
	M	2	5	3	
6	T	3	3	3	9
	P	3	3	3	
	M	2	2	5	
7*	T	1	2	7	10
	P	0	3	7	
	M	1	1	8	
8	T*	0	3	7	10
	P	0	6	4	
	M*	1	2	7	
9	T*	0	1	9	10
	P	1	3	6	
	M	1	4	5	
10	T	3	7	0	10
	P	3	7	0	
	M	2	6	2	

APPENDIX IV-4--Continued

		Rating			
Judge #	Theory Practice Methods*	Above Average	Average	Below Average	Total
11	T*	1	0	9	10
	P	3	7	0	
	M*	0	3	7	
12	T	6	0	4	10
	P	1	5	4	
	M	5	1	4	
13	T	1	3	6	10
	P	2	4	4	
	M	2	2	6	
14	T	4	3	3	10
	P	4	2	4	
	M	3	5	2	
15	T	2	4	4	10
	P	4	3	3	
	M	3	4	3	
16	T	2	4	4	10
	P	3	3	4	
	M	3	5	2	
17	T*	2	1	7	10
	P	3	3	4	
	M	4	4	2	
18	T	3	2	5	10
	P	3	4	3	
	M	2	2	6	
19	T	3	2	5	10
	P	3	4	3	
	M	2	3	5	
20	T	2	3	5	10
	P	2	3	5	
	M	1	7	2	
21	T	2	3	5	10
	P	5	3	2	
	M	2	5	3	

APPENDIX IV-4--Continued

Judge #	Theory Practice Methods*	Rating Above Average	 Average	 Below Average	Total
22	T	3	4	3	10
	P	4	3	3	
	M	4	1	5	
23	T	4	3	3	10
	P	5	4	1	
	M^t	6	3	1	
24	T	6	4	0	10
	P	2	7	1	
	M	3	4	3	
26	T^t	6	4	0	10
	P	6	1	3	
	M	6	2	2	
27	T^t	7	3	0	10
	P^t	6	4	0	
	M	63	3	1	
28	T	2	3	5	10
	P	2	4	4	
	M	1	3	6	
29	T*	1	1	9	11
	P*	2	2	7	
	M	1	5	5	
30	T	2	2	6	10
	P	2	3	5	
	M	2	3	5	
31	T	0	5	5	10
	P	1	6	3	
	M	1	3	6	
32	T	5	2	3	10
	P	5	2	3	
	M	6	2	2	
33	T	6	1	2	9
	P	3	3	3	
	M	3	3	3	

APPENDIX IV-4--Continued

		Rating			
Judge #	Theory Practice Methods*	Above Average	Average	Below Average	Total
34	T	5	1	4	10
	P	3	2	5	
	M	5	0	5	
35	T	4	1	5	10
	P	5	0	5	
	M	5	0	5	
36	T*	0	3	7	10
	P*	1	1	8	
	M	2	2	4	
37	T	3	4	3	10
	P	3	5	2	
	M	5	3	2	
38	T	4	2	4	10
	P	4	2	4	
	M	5	1	4	
39	T	5	1	4	10
	P	4	4	2	
	M	4	3	3	
40	T	3	3	4	10
	P	3	6	1	
	M	5	4	1	

APPENDIX VI-1

SUPPLEMENTARY TABLES FOR CHAPTER VI

TABLE 1

INDEX OF PRACTICE-ORIENTED SOCIALIZATION BY TYPE OF DOCTORATE

Amount of Practice-Oriented Socialization	Type of Doctorate	
	Behavioral Sciences	Education
Low	83%	17
Medium	15	47
High	2	36
	100%	100%
	(257)	(500)

TABLE 2

TYPES OF GRADUATE RESEARCH EXPERIENCE BY TYPE OF DOCTORATE

Types of Graduate Research Experience	Type of Doctorate	
	Behavioral Sciences	Education
Both worked in center and as research assistant	26%	21
Worked as an assistant but not in a center	39	31
Worked in a center but not as an assistant	4	2
Neither	17	23
No answer to one question or the other	13	23
	99%	100%
	(262)	(506)

TABLE 3

CAREER RESEARCH EXPERIENCE BY TYPE OF DOCTORATE

Amount of Career Research Experience	Type of Doctorate	
	Behavioral Sciences	Education
High	40%	23
Medium	31	29
Low	29	48
	100%	100%
	(232)	(430)

APPENDIX VI-2

MATRIX OF CORRELATIONS OF ITEMS IN INDEX OF PRACTICE-ORIENTED SOCIALIZATION

(Phi Coefficients)

	One or More Professional Education Courses	One or More Years of Elementary or Secondary School Teaching Experience
One or more years of school teaching experience	.551	-
Earned Ed.D.	.316	.367

INDEX OF PRACTICE-ORIENTED SOCIALIZATION

A study of the authors showed that three factors seemed to reflect their exposure to the norms and activities of educational practice. These factors were elementary or secondary school teaching experience, professional education courses, and name of degree. We formed the index of practice-oriented socialization by combining the answers from the following questions:

		Score	Frequency	Per Cent
"How many years have you taught in elementary or secondary school?"				
	None	0	382	42%
	One or more	1	511	57
	No answer	-	8	1
			901	100%
"How many graduate semesters, trimesters, or quarters did you take in . . . professional education [e.g., educational administration, higher education, guidance and counseling, curriculum and teaching in elementary or secondary school, subject matter courses (art, business, language, math, music, science, speech education, etc.)]?"				
	None	0	213	24%
	One or more	1	668	74
	No answer	-	20	2
			901	100%

	Score	Frequency	Per Cent
"Name of Degree"			
Ph.D.	0	567	63%
Ed.D.	1	219	24
Other degrees	-	6	1
No doctorate	-	102	11
No answer	-	7	1
		901	100%

Classification on the Index

	Score	Frequency	Per Cent
Low Practice-oriented Socialization	0	232	26%
Medium Practice-oriented Socialization	1	315	36
High Practice-oriented Socialization	2 or 3	334	38
		881	100%

APPENDIX VI-3

INDEX OF DEPARTMENTAL PRESTIGE

We classified authors according to their substantive specialty at a particular time (e.g., during an author's doctoral study, when he wrote his research paper, or when he returned the questionnaire). The following question was used to classify specialty at the time the research paper was written:

1.10 At the time of this research,

A. What was your main institutional affiliation?

1[ ] University
(a)____________ ____________
(name) (city and state)
2[ ] College

(b) What school or department were you in?
School or Department of Education
1[ ] undergraduate
2[ ] graduate
3[ ] joint undergraduate/graduate
Liberal arts department (please specify discipline)____________
4[ ] undergraduate
5[ ] graduate
6[ ] joint undergraduate/graduate
Other school or department (please specify)
____
7[ ] undergraduate
8[ ] graduate
9[ ] joint undergraduate/graduate

3[ ] School System ____________ ____________
(name) (city and state)

4[ ] State Department of Education ____________
(state)

5[ ] Private Agency ____________ ____________
(name) (city and state)

6[ ] Other (please specify) ____________ ____________
(name) (city and state)

The same question, with a slight change in wording to reflect the difference in time, was used to classify substantive specialty when the questionnaire was answered (see Question 2.2 in Questionnaire, Appendix II-5).

Specialty at the time of doctoral study was ascertained by combining the answers to two items in the questionnaire. Authors who indicated that their doctorate was in education (Question 3.1 in Questionnaire) were considered to be in schools of education. The doctoral departmental affiliation of behavioral scientists was determined by their major field (also part of Question 3.1 in Questionnaire).

At each of these three points in time we asked respondents for the name of their university. On the basis of the respondents' designation of their specialty and of their university, we classified their departmental affiliation as being more or less prestigious, using the evaluations of departmental prestige reported by Sieber and Cartter.

The Deans of Schools of Education in Sieber's study[1] mentioned the following graduate education departments as doing the best research:

	Number of Times Mentioned
Stanford	23
Wisconsin	18
Chicago	18
Harvard	17

---

[1] The Organization of Educational Research, Appendix C-2, Table 2).

	Number of Times Mentioned
Illinois	16
Teachers College, Columbia	14
Minnesota	11
University of Michigan	10
Ohio State	8
University of California, Berkeley	6
Oklahoma State	4
Syracuse	3
Boston University	2
State University of Iowa	2
Pittsburgh	2
Michigan State	2
Ball State	1
Indiana	1
University of Pennsylvania	1
Southern California	1
Texas	1
Washington University	1
Number of deans and coordinators replying:	(46)

Departmental Prestige of Education Doctorates	Frequency	Per Cent
Prestigious education department	262	54%
Less prestigious education department	226	46
	488	100%

Cartter[1] indicates that the following psychology graduate departments were considered as very good or outstanding (that is, they received one or two stars in his report);

- Brown
- University of California, Berkeley
- University of California, Los Angeles
- University of Chicago
- Columbia
- Cornell
- Duke
- Harvard
- Illinois
- Indiana
- Iowa
- Johns Hopkins
- Michigan
- Minnesota
- Northwestern
- Pennsylvania
- Stanford
- Wisconsin
- Yale

Among psychology doctorates, 47 per cent (78 authors) received doctorates from one of the above prestigious departments.

Cartter[2] reports that the following sociology graduate departments were considered as very good or outstanding (i.e., they received one or two stars in his report):

- University of California, Berkeley
- University of California, Los Angeles
- Chicago
- Columbia
- Cornell
- Harvard
- Johns Hopkins
- Michigan
- Minnesota
- North Carolina, Chapel Hill
- Northwestern
- Princeton
- Stanford
- Wisconsin

Among sociology doctorates, 63 per cent (26 authors) received degrees from the above prestigious departments.

---

[1] Cartter, *Assessment of Quality in Graduate Education*.

[2] *Ibid*.

APPENDIX VI-4

MATRIX OF CORRELATIONS OF ITEMS IN CAREER RESEARCH EXPERIENCE INDEX

(Phi Coefficients)

	Author spent 6 months or more when research was his primary activity	Author spent more than 50 per cent of his time on research while writing the paper
Author was a staff member of a research organization when he wrote his paper	.377	.341
Author spent more than 50 per cent of his time on research while writing the paper	.357	-

INDEX OF CAREER RESEARCH EXPERIENCE

Career research experience was measured by combining the answers to the following questions:

	Score	Frequency	Per Cent
"At the time of this research, how did you divide your professional time? (Please give your best estimate of the percent of time you spent on each activity.)"			
Per cent of time on research			
51 per cent or more	1	189	24
50 per cent or less	0	600	76
		789	100%
"Other than while you worked on your master's or doctoral thesis, has there ever been a period of six months or more during which research was your primary activity?"			
Yes	1	420	47%
No	0	472	53
		892	100%
"At that time (when the paper was written) were you a staff member of a research organization or bureau?"			
Yes	1	203	23%
No	0	675	77
		878	100%

This distribution of respondents on the index of career research experience is as follows:

Classification	Score	Frequency	Per Cent
High career research experience	2 or 3	225	29
Medium career research experience	1	219	29
Low career research experience	0	323	42
		767	100%

VALIDATION OF THE CAREER RESEARCH EXPERIENCE INDEX--CAREER RESEARCH EXPERIENCE BY AVERAGE RANK OF RESEARCH IN AUTHOR'S CAREER (UNWEIGHTED BY THE NUMBER OF YEARS HE HELD EACH POSITION)

Career Research Experience Index	Average Rank of Research in Author's Career		
	Always First 1.00 to 2.49	2.50-3.49	3.50 - No Research
High	70%	28	13
Medium	22	39	21
Low	9	33	66
	101%	100%	100%
	(135)	(320)	(254)

Tau Beta = .431

The high relationship between the index of career research experience and the average rank of research in the author's career provides an independent validation of the index of career research experience.

APPENDIX VII-1

SUPPLEMENTARY TABLES FOR CHAPTER VII

TABLES 1 and 2

ATTRACTION OF RESEARCH TOPIC BY SEX

	Sex	
	Female	Male
Attracted to research topic because of its importance for solving a concrete problem in education	25%	31
Attracted to topic because of its importance for testing theory	27	20
	(114)	(678)

TABLE 3

SOURCES OF RESEARCH TOPIC BY RELIGION

	Religion				
	Jewish	Areligious	Protestant	Catholic	Other
Most important source of topic was author's own reflection	21% (90)	18 (169)	13 (374)	13 (103)	12 (65)

TABLE 4

TWO POSSIBLE ATTRACTIONS OF RESEARCH TOPIC BY RELIGION

"What was the single most important feature of this topic which attracted you?"	Religion				
	Jewish	Areligious	Other	Protestant	Catholic
Important for developing or testing a behavioral science theory	24%	20	10	8	6
Important for solving a concrete problem in education	26	22	34	33	32
	(88)	(152)	(61)	(348)	(101)

APPENDIX VII-2

INTELLECTUAL ORIENTATION BY RELIGION

Intellectual Orientation (i.e., number of points given curiosity about substantive questions in a behavioral science field)	Religion				
	Jewish	Areligious	Protestant	Catholic	Other
4-9 points	40%	41	28	29	32
1-3 points	50	47	54	48	54
None	11	12	19	23	15
	101%	100%	101%	100%	101%
	(76)	(152)	(290)	(83)	(54)

APPENDIX VIII-1

INTELLECTUAL ORIENTATION BY ORGANIZATIONAL AFFILIATION

Intellectual Orientation (i.e., the number of points given curiosity about substantive questions in a behavioral science field)	Organizational Affiliation			
	University	College	School System	Private Agency and Other
None	33%	37	43	33
1-3 points	41	39	35	39
4-9 points	26	24	22	28
	100%	100%	100%	100%
	(691)	(67)	(54)	(64)

APPENDIX IX-1

INDEX OF PRODUCTIVITY

To form the index of productivity, we asked authors:

> How many empirical research studies (articles, monographs, or books) have you published in addition to the one referred to on page one, and what was the date of your first publication?

By dividing the total number of research articles published by the number of years[1] since their first publication, we computed an average annual number of research articles published, which is our index of research productivity. We averaged the number of publications so that authors with long careers would not appear to be disproportionately more productive than those with shorter careers.

DISTRIBUTION OF AUTHORS ON THE INDEX OF PRODUCTIVITY

Average Annual Number of Research Articles Published	Frequency	Per Cent
Less than one	306	41%
One or more, but less than two	224	30
Two or more	213	29
	743	100%

---

[1] There are minor methodological problems with using either date of first publication or date of doctorate as the basis for averaging the number of publications. Since authors in education generally take a longer time to complete the doctorate than do behavioral science doctorates, we decided that date of first publication was a less misleading basis for averaging the number of publications than date of doctorate, particularly since we wanted to compare the productivity of behavioral scientists and education doctorates.

APPENDIX IX-2

RESEARCH QUALITY BY RANK BY AGE

Research Quality	Age								
	43 - 70			38 - 42.			37 or Less		
	Academic Rank								
	Prof.	Assoc. Prof.	Ass't Prof.	Prof.	Assoc. Prof.	Ass't Prof.	Prof.	Assoc. Prof.	Ass't Prof.
ABOVE AVERAGE contribution to theory	26% (50)	[13] (15)	[33] (12)	- (6)	[38] (16)	[33] (18)	- (2)	[50] (10)	50 (22)
ABOVE AVERAGE contribution to practice	30 (50)	[27] (15)	[17] (12)	- (6)	[25] (16)	[28] (18)	- (2)	[20] (10)	27 (22)
ABOVE AVERAGE use of research methods	24 (50)	[13] (15)	[8] (12)	- (6)	[31] (16)	[44] (18)	- (2)	[50] (10)	50 (22)

APPENDIX IX-3

INDEX OF MOBILITY IN DEPARTMENTAL PRESTIGE

Departmental prestige was measured at three points in time (at the time of doctoral study, when the paper was written, and when the questionnaire was completed). We have already described how we measured departmental prestige in Appendix VI-3. High prestige at each point in time was given a score of 1. Lesser prestige at each point in time was given a score of 2. When these three measures were cross-tabulated, authors could have the following combinations, which were classified as follows:

Combinations	Classification	Frequency	Per Cent
111 (high,high,high)	Stayed the same--prestigious	66	11%
121, 211, 221	Moved up	47	8
112, 212, 122	Moved down	171	28
222	Stayed the same--less prestigious	333	54
		617	101%

APPENDIX IX-4

INDEX OF BEING REWARDED

The following questions and indices were used to form the index of being rewarded:

Index of Productivity (see Appendix IX-1)

	Score	Frequency	Per Cent
An average of two or more publications per year since their first publication	1	213	29%
An average of less than two publications per year since their first one	0	530	71
		743	100%

"Have you ever applied for a research grant from the USOE, NIMH, Ford, Rockefeller, or Carnegie foundations, or any other such organization that supports research?

IF YES, have you ever received a research grant from such an organization?

	Score	Frequency	Per Cent
Yes	1	398	44%
No	0	173	19
Did not apply for a grant	-	330	37
		901	100%

	Score	Frequency	Per Cent
Index of change in academic rank			
Moved up	1	153	40%
Stayed same	0	199	53
Can't tell	-	27	7
		379	100%
Index of mobility in departmental prestige (see Appendix IX-3)			
Moved up	1	66	11%
Other (stayed the same or moved down)	0	551	89
		617	100%

If an author received one or more of the above rewards he was considered to have been rewarded. The sample is distributed as follows:

	Frequency	Per Cent
Have been rewarded	552	61%
Have not been rewarded (or no information)	349	39
	901	100%